



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07H 21/04, C07K 13/00, 15/28, C12N 5/00, 5/10, 15/00, 15/12, 15/85, 15/86, G01N 33/53	A1	(11) International Publication Number: WO 94/12521 (43) International Publication Date: 9 June 1994 (09.06.94)
(21) International Application Number: PCT/US93/11310 (22) International Filing Date: 19 November 1993 (19.11.93) (30) Priority Data: 07/979,156 20 November 1992 (20.11.92) US (71) Applicant: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 201 West Seventh Street, Austin, TX 78701-2981 (US). (72) Inventors: LEE, Wen-Hwa; 11718 Jarvis Drive, San Antonio, TX 78245 (US). SHAN, Bei; 14825 Omicron Drive, #501, San Antonio, TX 78245 (US). (74) Agents: KONSKI, Antoinette, F. et al.; Campbell & Flores, 4370 La Jolla Village Drive, Suite 700, San Diego, CA 92122 (US).		(81) Designated States: CA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: CELLULAR GENES ENCODING RETINOBLASTOMA-ASSOCIATED PROTEINS		
(57) Abstract <p>This invention provides an isolated nucleic acid molecule encoding a retinoblastoma-associated protein, and isolated proteins having transcriptional factor E2F biological activity and RB-binding activity. This invention also provides vectors comprising an isolated nucleic acid molecule encoding a retinoblastoma-associated protein, mammalian cells comprising such vectors, antibodies directed to the retinoblastoma-associated protein and hybridoma lines producing monoclonal antibodies to such protein. This invention further provides methods for using such antibodies diagnostically and prognostically.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

CELLULAR GENES ENCODING RETINOBLASTOMA-ASSOCIATED
PROTEINS

This invention was made in part with Government support under grants issued by the National Institutes of Health Grant No. EY 05758 and Council for Tobacco Research to WHL. The Government may have certain rights in this invention.

FIELD OF THE INVENTION

This invention relates to the molecular cloning of cellular genes encoding retinoblastoma-associated proteins. In a more specific aspect it relates to the identification of a gene with properties of the transcription factor E2F.

Throughout this application various publications are referenced by partial citations within parentheses. The disclosures of these publications in their entireties are hereby incorporated by reference in this application in order to more fully describe the state of the art to which this invention pertains.

BACKGROUND OF THE INVENTION

The retinoblastoma gene (RB), the first tumor suppressor gene identified, encodes a nuclear phosphoprotein which is ubiquitously expressed in vertebrates (Friend, et al., Nature (London) 323:643-646 (1986); Lee, et al., Nature 329:642-645 (1987b); Fung, et al., Science 236:1657-1661 (1987)). Mutations of this gene which lead to inactivation of its normal function have been found not only in 100% of retinoblastomas but also in many other adult cancers including small cell lung-carcinoma (Harbour, et al., Science 241:353-357 (1988); Yokota, et al., Oncogene 3:471-475 (1988)), osteosarcoma (Toguchida, et al., Cancer Res. 48:3939-3943 (1988)), bladder carcinoma (Horowitz, et al., Science 243:937-940 (1989)), prostate

carcinoma (Bookstein, et al., PNAS USA 87:7762-7766 (1990a)) and breast cancer (Lee et al., Science 241:218-221 (1988)). Reconstitution of a variety of RB-deficient tumor cells with wild-type RB leads to suppression of their
5 neoplastic phenotypes including their ability to form tumors in nude mice (Huang, et al., Science 242:1563-1566 (1988); Sumegi, et al., Cell Growth Diff. 1:247-250 (1990); Bookstein, et al., Science 247:712-715 (1990b); Goodrich, et al., Can. Res. 52:1968-1973 (1992); Takahashi, et al.,
10 PNAS USA 88:5257-5261 (1991); Chen, et al., Cell Growth Diff. 3:119-125 (1992)). These results provide direct evidence that RB protein is an authentic tumor suppressor.

RB performs its function at the early G1/G0 phase of the cell cycle as substantiated by several observations:
15 first, the phosphorylation of RB, presumably by members of the Cdk kinase family (Lin, et al., EMBO J. 10:857-864 (1991); Lee, et al., Cell Cycle. 61:211-217 (1991)), fluctuates with the cell cycle (Chen, et al., Cell 58:1193-1198 (1989); Buchkovich, et al., Cell 58:1097-1105 (1989);
20 DeCaprio, et al., Cell 58:1085-1095 (1989)); second, the unphosphorylated form of RB is present predominantly in the G0/G1 stage (Chen, et al., 1989, supra.; DeCaprio. et al., 1989, supra.); third, microinjection of the unphosphorylated RB into cells at early G1 phase inhibits
25 their progression into S phase (Goodrich, et al., Cell 67:293-302 (1991)). These observations suggest that RB may serve as a critical regulator of entry into cell cycle and its inactivation in normal cells could lead to deregulated growth.

30 How RB functions is the subject of intense inquiry. Two known biochemical properties of the RB protein have been described; one is its intrinsic DNA binding activity which was mapped to its C-terminal 300 amino acid residues (Lee et al., 1987b, supra.; Wang, et
35 al., Cell Growth Diff. 1:429-437 (1990b)); another is its

ability to interact with several oncoproteins of the DNA tumor viruses (DeCaprio, et al., Cell 54:275-283 (1988); Whyte, et al., Nature 334:124-129 (1988); Dyson, et al., Science 243:934-937 (1989)). This interaction was mapped
5 to two discontinuous regions at amino acids 379-545 and 575-678, designated as the T-binding domains (Hu, et al., EMBO J. 9:1147-1155 (1990); Huang, et al., EMBO J. 9:1815-1822 (1990)). Interestingly, mutations of the RB proteins in tumors were frequently located in these same regions
10 (Bookstein and Lee, CRC Crit. Rev. Oncogenesis 2:211-227 (1991)). These results imply that the T-binding domains of RB proteins are functionally important and the interaction of RB with these oncoproteins may have profound biological significance. The identification of cellular proteins that
15 mimic the binding of T to RB revealed a potentially complicated network. Several proteins including c-myc (Rustgi, et al., Nature 352:541-544 (1991)), Rb-p1, p2 (Defeo-Jones, et al., Nature 352:251-254 (1991)) and 8-10 other proteins (Kaelin, et al., Cell 64:521-532 (1991));
20 Lee, et al., 1991, supra.; Huang, et al., Nature 350:160-162 (1991)) have been shown to bind to RB in vitro.

As the foregoing demonstrates, there clearly exists a pressing need to identify and characterize the cellular affiliates of the retinoblastoma gene. The
25 present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

This invention provides an isolated nucleic acid molecule encoding a retinoblastoma-associated protein, and
30 isolated proteins having transcriptional factor E2F biological activity and RB-binding activity.

This invention further provides vectors such as plasmids and viruses comprising a DNA molecule encoding a

retinoblastoma-associated protein adapted for expression in a bacterial cell, a yeast cell, or a mammalian cell.

This invention provides a mammalian cell comprising a DNA molecule encoding a retinoblastoma-associated protein.

This invention provides an antibody capable of specifically binding to a retinoblastoma-associated protein. This invention also provides hybridoma cell lines that produce monoclonal antibodies and methods of using these antibodies diagnostically and prognostically.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the results of RB-sandwich screening. λ gt11 cDNA expression libraries were plated and screened using the RB-sandwich containing purified p56-RB, anti-RB antibody, and alkaline-phosphatase conjugated secondary antibody. A and B, a diagram of the RB-sandwich screening. C and D, hybridized filters with the RB-sandwich (left halves of the filters) in which the positive signal indicates a RbAp-RB complex (C) or T-antigen-RB complex (D). The right halves of the filters were probed with the RB-minus sandwich.

Figure 2 shows binding of RbAps to RB in vitro. The cDNA insert from each clone (Ap4, 6, 9, 10, 11, 12, 15) was subcloned into the pFLAG plasmid and the lysates of FLAG-Ap fusion proteins were mixed with the GST-RB beads (R) or GST beads alone (C). The bound proteins were then analyzed by immunoblot using a monoclonal anti-FLAG antibody. The arrows indicate the FLAG-Aps bound to the GST-RB beads, which were detected by the anti-FLAG antibody. BAP = FLAG bacterial alkaline phosphatase fusion protein.

Figure 3 shows cell cycle dependent expression of Apl2. Total RNA from CV1 cells synchronized at various stages of the cell cycle was denatured and analyzed by formaldehyde gel electrophoresis. The RNA blot was hybridized with a ³²P-labeled Apl2 cDNA insert (G12). Lane 1, early G1; lane 2, G1/S boundary; lane 3, S phase (4 hours after aphidicolin release); lane 4, S phase (18 hours after replating starved cells); lane 5, M phase. The size of the mRNA (designated by an arrow) was determined by migration of the rRNA 28S and 18S, which were run on a parallel lane next to the RNA samples.

Figure 4 shows the restriction map and nucleic and amino acid sequences of Apl2. Clone A6, 2,492 nucleotides, was completely sequenced (SEQ ID NOS: 13-14). A: restriction map of Apl2 (A6) which has the longest open reading frame. G12 is the original Apl2 clone obtained by the RB-sandwich screening. A6 and B6 were isolated by rescreening of cDNA libraries. Only restriction sites used in the construction of Apl2 derivatives are shown. B: sequence of Apl2 and predicted amino acid sequence. The squares indicate the leucine repeats. Two putative Cdk phosphorylation sites are underlined.

Figure 5 shows that Apl2 binds specifically to the hypophosphorylated form of RB at regions similar to T. A, Lane 1: a Molt4 lysate immunoprecipitated using a monoclonal anti-RB antibody, mAb11D7. Lane 2: molecular marker. Lanes 3-5: Molt4 cell lysates (5x10⁶ cells) were mixed with GST beads (lane 3), GST-Apl2 (lane 4) and GST-T (lane 5) beads. After washing, the RB bound to the GST fusions was analyzed by immunoblotting using a monoclonal anti-RB antibody, mAb11D7. B: a panel of RB mutant proteins expressed in a bacterial pET-T7 expression system. The T-binding domains are highlighted. C-D: the bacterially expressed wild type (pETRbc) or mutant RB proteins (pETB2, Ssp, Xs, M8, M6, M9, Nm) were mixed with

the GST-Ap12 (C) or GST-T (D) beads and the bound proteins were measured by Western blot analysis using a monoclonal anti-RB antibody, mAb245.

Figure 6 shows that the C-terminal region of Ap12 is required for RB-binding. A series of GST-Ap12 derivatives, P3, SH5, XH9, SX4, and XX4 were constructed (shown in panel B) and used for RB binding. The bacterially expressed pETRbc (wild type RB) was mixed with the GST-Ap12 beads and analyzed by Western blot analysis using a monoclonal anti-RB antibody, mAb245. The polypeptide encoding region for P3 is amino acids 362-476; SH5, aa 162-476; XH9, aa 1-476; SX4, aa 162-455; XX4, aa 1-455. The arrow indicates the position of p110-RB.

Figure 7 shows that Ap12 binds specifically to the E2F recognition sequence. The lysates prepared from the bacterially expressed derivatives of GST-Ap12 (P3, SH5, XH9) and GST-Ap9, GST-Ap15 and GST alone were used for DNA mobility shift assays. The probe was a DNA fragment containing two E2F recognition sites, which was ³²P-end-labeled by Klenow fill-in reaction. A: GST-Ap12SH5 binds to the E2F-specific sequence. As a positive control, a partially purified E2F protein from HeLa cells was also used. DNA fragments containing either the wild type E2F sites or mutated E2F sites were used as competitors. Lane 1: probe alone; Lane 2: E2F + probe; Lane 3: E2F + probe + wt competitor; Lane 4: E2F + probe + mutant competitor; Lane 5: SH5 + probe; Lane 6: SH5 + probe + wt competitor; Lane 7: SH5 + probe + mutant competitor. B: RB interacts with the Ap12-E2F DNA complex. Lane 1: probe alone; Lane 2: SH5 + probe; Lane 3: SH5 + p56-RB (0.25 µg), incubate for 15 minutes, followed by probe addition; Lane 4: p56-RB + probe; Lane 5: SH5 + probe; Lane 6: SH5 + probe for 15 minutes, then p56-RB was added. C: DNA binding domain of Ap12 is located at a region containing a potential bZIP motif. Lane 1: P3, 200ng; Lane 2: P3, 400ng; Lane 3,

SH5, 20ng; Lane 4: SH5, 40ng; Lane 5: XH9, 20ng; Lane 6: XH9, 40ng; Lane 7: GST alone, 200ng; Lane 8: GST-Ap9, 200 ng; Lane 9: GST-Ap15, 200ng.

Figure 8 shows that the C-terminus of Ap12 serves
5 as an activation domain when fused to the GAL4 DNA binding domain in yeast. Fusion proteins of GAL4 (amino acids 1-147) and either G12 (AP12, amino acids 362-476), 12B6 (AP12, amino acids 22-476) or Rb2 (RB, amino acids 301-928) were expressed in yeast as detailed below. Plasmids were
10 used to transform Y153 to tryptophan prototrophy, and single colonies of each transformation were streaked on dropout media lacking tryptophan. Following 1 day of growth at 30°C, cells were analyzed for β -galactosidase activity using a colony lift assay.

Figure 9 shows that Ap12 transactivates a
15 promoter with E2F recognition sites. A: a diagram of the Ap12 cDNA expression vectors. PA, poly(A). B: transcriptional activation of a promoter with E2F recognition sequences. 10 μ g of either pA₁₀CAT or pE2FA₁₀CAT
20 was cotransfected with 10 μ g of CMV-Ap12-Stu or CMV-Ap12-RH into monkey kidney CV1 cells. The cells were harvested after 48 hours and CAT activities were measured. CMV-E4 was cotransfected with the reporter plasmids as well as the reporter plasmids alone to serve as a control.

Figure 10 shows the partial nucleic acid sequence
25 of clone Ap2. p = 5' sequence (SEQ ID NO: 5); r = 3' sequence (SEQ ID NO: 6).

Figure 11 shows the partial nucleic acid sequence
of clone Ap8. p = 5' sequence (SEQ ID NO: 7); r = 3'
30 s quence (SEQ ID NO: 8).

Figure 12 shows the partial nucleic acid sequence
of clone Ap15. p = 5' sequence (SEQ ID NO: 9); r = 3' sequence (SEQ ID NO: 10).

Figure 13 shows the full length nucleic acid sequence of clone Ap4 (SEQ ID NO: 11).

Figure 14 shows the full length nucleic acid sequence of clone Ap10 (SEQ ID NO: 12).

5

DETAILED DESCRIPTION OF THE INVENTION

The retinoblastoma protein interacts with a number of cellular proteins to form complexes which can be crucial for its normal physiological function. To identify these proteins, nine distinct gene cDNAs were cloned by
10 direct screening of cDNA expression libraries using purified RB protein as a probe. Preliminary characterization of these clones indicates that a majority of these genes encode novel proteins. One of them, Ap12, expresses a 2.8 Kb mRNA in a cell cycle-dependent manner.

15

The longest cDNA isolate of Ap12 encodes a putative protein of 476 amino acids with several features characteristic of transcription factors. The C-terminal 114 amino acids of Ap12 binds to unphosphorylated RB in regions similar to where T-antigen binds and has
20 transactivation activity. A region near the N-terminus contains a putative leucine zipper flanked by basic residues and is capable of specifically binding to an E2F cognate sequence. Expression of Ap12 in monkey kidney CV1 cells significantly enhanced E2F-dependent transcriptional
25 activity. Although the E2F gene has not been cloned and its identity is based solely on the ability to recognize and bind to a specific DNA sequence, these results establish that the novel clones encode proteins with known properties of the transcription factor E2F and which bind
30 RB.

Accordingly, the present invention provides an isolated nucleic acid molecule encoding a retinoblastoma-

associated protein. As used herein, the term "isolated nucleic acid molecule" refers to a nucleic acid molecule that is in a form that does not occur in nature. One means of isolating a human retinoblastoma nucleic acid molecule is to probe a human cDNA expression library with a natural or artificially designed antibody to retinoblastoma, using methods well known in the art (see Sambrook et al. Molecular Cloning: A Laboratory Manual 2d ed. (Cold Spring Harbor Laboratory 1989)) which is incorporated herein by reference). DNA and cDNA molecules which encode human retinoblastoma-associated polypeptides can be used to obtain complementary genomic DNA, cDNA or RNA from human, mammalian or other animal sources. The isolated nucleic acids can also be used to screen cDNA libraries to isolate other genes encoding RB-associated proteins.

The present invention provides soluble retinoblastoma-associated polypeptides that have DNA binding and RB binding activity. For the purposes of illustration only, nucleic acid sequences encoding the polypeptides are identified in Figures 4 and 10-14. The nucleic acid sequences encoding the soluble retinoblastoma-associated polypeptide are included within the sequences set forth in Figures 4 and 10-14.

As used herein "retinoblastoma-associated polypeptide" means a polypeptide having that has DNA binding as well as an RB-binding activity. Examples of retinoblastoma-associated polypeptides substantially the same as the amino acid sequence of clone Ap12, shown in Figure 4, or the amino acid sequence encoded by the nucleic acid sequences of clones Ap 2, 4, 8, 10 and 15, or active fragments thereof. As used herein, "an active fragment or biologically-active fragment" refers to any portion of the retinoblastoma-associated polypeptide shown in Figure 4, or that encoded by clones Ap 2, 4, 8, 10 and 15 shown in Figures 10-14. Methods of determining whether a

polypeptide can bind RB are well known to those of skill in the art, for example, as set forth herein.

As used herein, the term "purified" means that the molecule or compound is substantially free of
5 contaminants normally associated with a native or natural environment. The purified polypeptides disclosed herein include soluble polypeptides. For example, the purified soluble polypeptide can be obtained from a number of methods. The methods available for the purification of
10 proteins include precipitation, gel filtration, ion-exchange, reversed-phase, and affinity chromatography. Other well-known methods are described in Deutscher et al., Guide to Protein Purification: Methods in Enzymology Vol. 182, (Academic Press 1990), which is incorporated herein by
15 reference. Alternatively, a purified polypeptide of the present invention can also be obtained by well-known recombinant methods as described, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual 2d. ed. (Cold Spring Harbor Laboratory 1989), also incorporated
20 herein by reference. An example of this means for preparing soluble retinoblastoma-associated polypeptide is to express nucleic acid encoding the retinoblastoma-associated polypeptide in a suitable host cell, such as a bacterial, yeast or mammalian cell, using methods well
25 known in the art, and recovering the expressed soluble protein, again using methods well known in the art. The soluble polypeptide and biologically active fragments thereof can also be produced by chemical synthesis. Synthetic polypeptides can be produced using Applied
30 Biosystems, Inc. Model 430A or 431A automatic polypeptide synthesizer and chemistry provided by the manufacturer. The soluble polypeptide can also be isolated directly from cells which have been transformed with the expression v ctors described below in more detail.

The invention also encompasses nucleic acid molecules which differ from that of the nucleic acid molecules shown in Figures, but which produce the same phenotypic effect. These alter d, but phenotypically equivalent nucleic acid molecules are referred to "equivalent nucleic acids." This invention also encompasses nucleic acid molecules characterized by changes in non-coding regions that do not alter the phenotype of the polypeptide produced therefrom when compared to the nucleic acid molecule described hereinabove. This invention further encompasses nucleic acid molecules which hybridize to the nucleic acid molecule of the subject invention. As used herein, the term "nucleic acid" encompasses RNA as well as single- and double-stranded DNA and cDNA. In addition, as used herein, the term "polypeptide" encompasses any naturally occurring allelic variant thereof as well as man-made recombinant forms.

The invention further provides the isolated nucleic acid molecule operatively linked to a promoter of RNA transcription, as well as other regulatory sequences. As used herein, the term "operatively linked" means positioned in such a manner that the promoter will direct the transcription of RNA off the nucleic acid molecule. Examples of such promoters are SP6, T4 and T7. Vectors which contain both a promoter and a cloning site into which an inserted piece of DNA is operatively linked to that promoter are well known in the art. Preferable, these vectors are capable of transcribing RNA in vitro or in vivo. Examples of such vectors are the pGEM series (Promega Biotech; Madison, WI).

This invention provides a vector comprising this isolated nucleic acid molecule encoding a retinoblastoma-associated polypeptide. Examples of vectors are viruses, such as bacteriophages, baculoviruses and retroviruses, cosmids, plasmids and other recombination vectors. Nucleic

acid molecules are inserted into vector genomes by methods well known in the art. For example, insert and vector DNA can both be exposed to a restriction enzyme to create complementary ends on both molecules that base pair with each other and which are then joined together with a ligase. Alternatively, synthetic nucleic acid linkers can be ligated to the insert DNA that correspond to a restriction site in the vector DNA, which is then digested with a restriction enzyme that recognizes a particular nucleotide sequence. Additionally, an oligonucleotide containing a termination codon and an appropriate restriction site can be ligated for insertion into a vector containing, for example, some or all of the following: a selectable marker gene, such as neomycin gene for selection of stable or transient transfectants in mammalian cells; enhancer/promoter sequences from the immediate early gene of human cytomegalovirus (CMV) for high levels of transcription; transcription termination and RNA processing signals from SV40 for mRNA stability; SV40 polyoma origins of replication and ColE1 for proper episomal replication; versatile multiple cloning sites; and T7 and SP6 RNA promoters for in vitro transcription of sense and anti-sense RNA. Other means are available and one well known for those of skill in the art.

Also provided are vectors comprising a DNA molecule encoding a human retinoblastoma-associated polypeptide, adapted for expression in a bacterial cell, a yeast cell, a mammalian cell and other animal cells. The vectors additionally comprise the regulatory elements necessary for expression of the DNA in the bacterial, yeast, mammalian or animal cells so located relative to the DNA encoding retinoblastoma-associated polypeptide as to permit expression thereof. Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression

vector includes a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG (Sambrook et al., supra.). Similarly, a eucaryotic expression vector includes a heterologous or
5 homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors can be obtained commercially or assembled by the sequences described in methods well known in the art, for
10 example the methods described above for constructing vectors in general. Expression vectors are useful to produce cells that express the polypeptide.

This invention provides a host cell, e.g. a mammalian cell, containing a nucleic acid molecule encoding
15 a human retinoblastoma-associated polypeptide. An example is a mammalian cell comprising a plasmid adapted for expression in a mammalian cell. The plasmid has a nucleic acid molecule encoding a retinoblastoma-associated polypeptide and the regulatory elements necessary for
20 expression of the polypeptide. Various mammalian cells may be utilized as hosts, including, for example, mouse fibroblast cell NIH3T3, CHO cells, HeLa cells, Ltk- cells, etc. Expression plasmids such as those described supra can be used to transfect mammalian cells by methods well known
25 in the art such as calcium phosphate precipitation, DEAE-dextran, electroporation or microinjection.

Also provided are antibodies having specific reactivity with the retinoblastoma-associated polypeptides of the subject invention, such as anti-Apl2 antibody, or
30 any antibody having specific reactivity to a retinoblastoma-associated polypeptide. Immunologically active fragments of antibodies are encompassed within the definition of "antibody." Identification of immunologically active fragments can be performed, for
35 example, as detailed below. The antibodies of the

invention can be produced by any method known in the art. For example, polyclonal and monoclonal antibodies can be produced by methods well known in the art, as described, for example, in Harlow and Lane, Antibodies: A Laboratory
5 Manual (Cold Spring Harbor Laboratory 1988), which is incorporated herein by reference. The polypeptide, particularly retinoblastoma-associated polypeptide of the present invention, can be used as the immunogen in generating such antibodies. Altered antibodies, such as
10 chimeric, humanized, CDR-grafted or bifunctional antibodies can also be produced by methods well known to those skilled in the art. Such antibodies can also be produced by hybridoma, chemical synthesis or recombinant methods described, for example, in Sambrook et al., supra,
15 incorporated herein by reference. The antibodies can be used for determining the presence or purification of the retinoblastoma-associated polypeptide of the present invention. With respect to the detecting of such polypeptides, the antibodies can be used for in vitro
20 diagnostic or in vivo imaging methods for diagnosing or prognosing pathologies associated with loss of functional RB protein.

Any of the above-identified novel compositions of matter may be combined with a pharmaceutically acceptable
25 carrier. As used herein, "pharmaceutically acceptable carrier" mean any of the standard carriers, such as saline, emulsion and various wetting agents. These compositions can be used for the preparation of medicaments for the treatment of pathologies associated with the loss of
30 functional RB protein.

Immunological procedures useful for in vitro detection of the target retinoblastoma-associated polypeptide in a sample include immunoassays that employ a detectable antibody. Such immunoassays include, for
35 example, ELISA, Pandex microfluorimetric assay,

agglutination assays, flow cytometry, serum diagnostic assays and immunohistochemical staining procedures which are well known in the art. An antibody can be made detectable by various means well known in the art. For example, a detectable marker can be directly or indirectly attached to the antibody. Useful markers include, for example, radionuclides, enzymes, fluorogens, chromogens and chemiluminescent labels.

Identification of RB-associated proteins. The simplest model for RB function is that relatively few target molecules which play central roles in cellular function are regulated by the retinoblastoma protein. Inactivation of RB by any one of three means, phosphorylation (Chen, et al., 1989, supra.; DeCaprio, et al., 1989, supra.), mutations (Shew, et al., PNAS USA 87:6-10 (1990)) or oncoprotein perturbation (DeCaprio, et al., 1988, supra.; Goodrich, et al., 1991, supra.; Whyte, et al., 1988, supra.), could potentially uncouple RB connections and lead to deregulated growth. Until this report, there were, indeed, only a limited number of molecules that were known to be capable of interacting with RB, such as two proteins of unknown function, p1 and p2, the myc protein and 8-10 other unidentified proteins. To genetically and biochemically dissect the RB network, it is essential to identify as many of the genes encoding interactive partners of RB as possible. To maximize the cloning probability, two different approaches were undertaken. One approach was to use a two-hybrid method developed by Field and his colleagues (Fields and Sung, Nature 340:245-246 (1989)) based on the yeast GAL4 system to select for protein-protein interaction in vivo. The other approach, described herein, was to use an RB-sandwich to screen λ gt11 cDNA expression libraries. The advantage of using this one-step RB-sandwich procedure is its simplicity, directness, and the clone isolated should encode a fusion protein that would directly interact with

RB in the absence of potential bridging proteins. Screening was performed using SV40 large T antigen as a positive control. A λ gt11 phage expressing T antigen was constructed for this purpose and the association between RB
5 and T can be readily detected by this method.

Using this approach, 9 clones were isolated. All the proteins encoded by these clones are located in the nucleus. This is an important criteria for any protein that could interact with RB in a biologically significant
10 manner, since the interaction probably would occur in the nucleus (Lee, et al., 1987b, supra.).

Transcription factors as targets of regulation by the RB protein. If the cellular function of RB is to restrict entry of cells into G1 (Goodrich, et al., 1991,
15 supra.), the genes important for G1 progression and entrance into S phase should be regulated directly or indirectly by RB. The transcription factor E2F is known to associate with RB in a cell-cycle-dependent manner (Mudryj, Cell 28:1243-1253 (1991); Shirodkar, Cell 68:157-166
20 (1992)), with a tight association being prevalent in the G0/G1 stage but not in S or M phases. There are several genes including myc, DHFR, and myb that may be subject to E2F transcriptional control (Hiebert, et al., PNAS USA 86:3594-3598 (1989); Mudryj, et al., EMBO J. 9:2179-2184
25 (1990)). It is reasonable to propose that RB sequesters E2F in the G0/G1 stage in an inactive conformation. Its release from the RB complex allows it to assume an active conformation that is capable of influencing its target genes through interactions with E2F DNA-binding sites and
30 the general transcriptional machinery. An important challenge is to determine the identity of the E2F target genes and to ascertain their role in the control of the cell cycle.

There is increasing evidence to support this simple model of RB function, which is now further supported by the finding that, in the collection of 9 newly cloned RB-associated proteins, one is a known eukaryotic upstream binding factor (UBF) which recognizes and binds to the ribosomal RNA promoter, and activates transcription mediated by RNA polymerase I through cooperative interactions with SL1 (Jantzen, et al., Nature 344:830-836 (1990)), and another, Apl2, has properties consistent with those proposed for the E2F transcription factor. The accumulation of Apl2 mRNA around six hours post stimulation with serum coincides with the pattern of expression of delayed-early growth response genes (Lau and Nathans, "Genes induced by serum growth factors" In The Hormonal control regulation of gene transcription, ed. P. Cohen & J.G. Foulkes, Elsevier Science Publishers, pp. 257-293 (1991)). The maximal level of Apl2 mRNA accumulates at the G1/S boundary, establishing that it has a role in controlling cells of entry into S phase. Also, the protein binds only to unphosphorylated RB at domains similar to those bound by T. Most interestingly, Apl2 recognizes the E2F cognate sequence and transactivates the promoter carrying such specific sequence.

Apl2 encodes a putative bZIP transcription factor. From the preliminary characterization of this gene, the putative protein deduced from the longest open reading frame is 476 amino acids in length although the initiating methionine has yet to be defined. The predicted molecular weight of the putative protein is about 51 kd which is close to the 60 kd protein immunoprecipitated by the anti-Apl2 antibody. The C-terminal region of Apl2 which binds to RB protein and has a transactivation activity, is very acidic, a hallmark of the transactivation domain of several known transcription factors such as GAL4 and VP16 (Sadowski, et al., Nature 335:563-564 (1988); Mitchell and Tjian, Science 245:371-378 (1989)). The DNA

binding domain appears to be located at the middle region of the protein which features a putative leucine zipper motif flanked by stretches of basic amino acids. Since Apl2 has most of the features that are characteristic of E2F, it can be considered to either encode E2F or a protein in the E2F family. Thus it is likely that E2F is also a bZIP protein which is intriguing since this is a class of transcription factors intimately involved in cell growth (e.g., fos and jun) and differentiation (e.g., C/EBP). Another hallmark of the bZIP family is a propensity to form a diverse array of heterodimeric associations among its members which adds a new layer of regulation to the control of E2F.

This vast array of possibilities presents an almost unlimited opportunity for the cell to intricately regulate the proteins involved in fine control of the cell cycle. The availability of the Apl2/E2F clone will facilitate the further elucidation of the connection between RB, E2F and cellular proliferation.

To identify the cellular affiliates of RB and to initiate the elucidation of the RB interactive cellular network, several approaches were taken to clone genes encoding RB-associated proteins. Described herein are the results from one of these approaches: screening of λ gt11 expression libraries using RB as a probe. Nine distinct genes were cloned, one of which, Apl2, has characteristics which suggest that it encodes the transcription factor E2F. Clones Ap 2, 4, 8, 10, 12 and 15 all encode RB-associated proteins and are all involved in cell cycle control.

Identification of RB-associated proteins (RbAps).
Two λ gt11 cDNA expression libraries were constructed and screened using the purified p56-RB protein (amino acids 376-928) which includes both T-binding domains and entire C-terminal region (Lee, et al., 1991, supra.) as probe.

This probe is referred to as a RB-sandwich since it contains RB protein, rabbit anti-RB antibody, (0.47) (Wang, et al., Cell Growth Diff. 1:233-239 (1990a)), and alkaline phosphatase conjugated goat anti-rabbit IgG. (see Materials and Methods). Figure 1 illustrates a diagram of the sandwich screening strategy (1A and 1B). Since the association of RB and SV40 T-antigen is well documented (DeCaprio, et al., 1988, supra.), a λ gt11 phage expressing T-antigen was constructed and screened using the RB-sandwich to serve as a positive control (shown in Figure 1-D). As an example (Figure 1-C), one of the clones' (Ap12) fusion product, was readily detected by this method. One half of each filter was used for binding to the RB-sandwich and the other half to the sandwich minus RB protein. The latter probe served as a control for the background binding due to any cross-reaction of the RB antibody or goat anti-rabbit antibody with bacterial proteins. After 5 rounds of screening of 1×10^6 recombinant phage, 12 clones emerged as candidate genes encoding RB-associated proteins. These clones are designated RbAp1, 2, 4, 6, 8, 9, 10, 11, 12, 13, 14, 15.

These 12 putative RbAp cDNAs were subcloned into the pGEM plasmid and a partial sequence of 500 to 600bp from each clone was obtained. A comparison with known gene sequences present in the GENBANK database, RbAp1, 2, 4, 8, 10, 12, 13, 14, 15 appear to be novel genes that contain no significant homology to any known genes. However, three clones matched previously identified genes: RbAp6 is identical to nuclear lamin C (McKeon et al., Nature 319:463-468 (1986); Fisher et al., PNAS USA 83:6450-6454 (1986)); RbAp9 encodes a product partially homologous to the β subunits of G protein (Gullemont et al., PNAS USA 86:4594-4598 (1989)); and RbAp11 codes for the upstream binding factor (UBF) that binds to the ribosomal RNA gene promoter (Jantzen, et al., supra.). Cross-hybridization and sequencing data showed that RbAp1, 10, 13, and 14 are

identical. Table 1 summarizes the preliminary characterization of all the cloned RbAps.

RbAp clones 2, 4, 8, 10, 12, and 15 are targets for RB, p110^{RB}, binding and all function in cell cycle control. It is possible that the retinoblastoma-associated proteins encoded by the RbAp clones are positive elements for cell proliferation. Rb binds to the protein products of these clones and, therefore, inhibits their proliferative function. As a result, the RbAp protein products cannot function positively and, therefore, are unable to promote cell cycle progression. Alterations in the RbAp ability to bind RB can result in an oncogenic effect. Assays detecting such alterations and/or mutations could determine malignancy and function as diagnostic tools for hyperproliferative diseases. Examples of hyperproliferative pathologies include, but are not limited to thyroid hyperplasia, psoriasis, Li-Fraumeni syndrome including breast cancer, sarcomas and other neoplasms, bladder cancer, colon cancer, lung cancer, benign prostatic hypertrophy and various leukemias and lymphomas. The present invention also provides antagonists of such altered and/or mutated RbAps for use in therapeutics for cancer and other hyperproliferative pathologies.

Table 1. Initial characterization of RB-associated proteins. The size of cDNA of each clone was determined by the EtBr staining of the agarose gel after digestion of the phage DNA with EcoRI. The size of mRNAs was measured by the RNA blot analysis using 28s and 18s rRNA as markers. The partial sequence from each clone was used to search GENBANK database to determine the identity of the clones. The nuclear localization was determined by immunostaining and cell fractionation (data not shown). nd = not determined.

RbAp	Length of cDNA (kb)	Size of mRNA (kb)	<u>in vitro</u> Binding	Identity	Subcellular Localization
1,10,13,14	2.8	7.1	+	Novel	Nucleus
2	1.6	3.6	nd	Novel	nd
4	1.7	6.7	+	Novel	Nucleus
6	1.5	2.1	+	Lamin C	Nucleus
8	1.8	6.9	nd	Novel	nd
9	0.7	1.3	+	GB-like	Nucleus & Membrane
11	1.5	3.2	+	UBF	Nucleus
12	1.4	2.8	+	Novel	nd
15	1.5	6.5	+	Novel	Nucleus

Binding of RbAps to RB in vitro. To confirm the association of RB protein with RbAps, the cloned cDNA inserts were subcloned into the plasmid pFLAG (IBI). This plasmid is designed for expressing Flag-fusion proteins in bacteria which can then be detected using an antibody against the Flag segment of the fusion. To facilitate the binding assay, the p56-RB was fused with the glutathione S-transferase (Gst) gene, expressed and purified by glutathione agarose chromatography (Gst-RB) (Smith and Johnson, Gene 67:31-40 (1988)). To perform the RB-binding assay, the FLAG-Ap lysates were mixed with the Gst-RB or Gst beads alone (no RB). As an additional negative

control, FLAG-BAP (bacterial alkaline phosphatase) was also mixed with the Gst and Gst-RB beads. After extensive washing, the bound fusion proteins were eluted and analyzed by Western blotting using the anti-FLAG monoclonal
5 antibody. The results demonstrate that all RbAps examined are able to bind to the Gst-RB beads but not to the control Gst beads (Figure 2). Among these clones, the binding affinity varied from Ap15, the weakest, to Ap12, the strongest.

10 The level of Ap12 mRNA is regulated during the cell cycle. Since Ap12 consistently showed the strongest binding signal during screening, it was selected for further study. The clone has an insert of 1.4 kb with a
15 about 1.0 kb untranslated region and an open reading frame of 114 amino acids. RNA blot analysis was performed to determine the size of the mRNA and its pattern of expression during cell cycle progression. Normal monkey
20 kidney CV1 cells were plated in fresh medium with 10% serum in the presence of Lovastatin for 36 hours (to arrest the cell in G1 phase) (Jakobisiak, et al., PNAS USA 88:3628-3632 (1991); Keyomarsi, et al., Can. Res. 51:3602-3609 (1991)) or aphidicolin (10 µg/ml) for 16 hours (to arrest the cells at the G1/S boundary), then released for 4 hours (to synchronize the cells in S phase) or incubated in the
25 presence of nacodazole for another 16 hours (to allow the cells to progress to M phase) (Goodrich, et al., 1991, supra.). Total RNA from each stage was prepared for blot analysis using the Ap12 cDNA as a probe. A 2.8 kb mRNA was detected at the G1/S boundary and in S phase, but was
30 undetectable in early G1 or M phase (Figure 3). As a control, the expression pattern of Ap9 does not change during the cell cycle. Consistent with this observation, an increase of Ap12 mRNA expression was observed between 2 and 6 hours after serum stimulation. These findings
35 establish that Ap12 can be involved in cell cycle progression.

Sequence analysis of Apl2. It is apparent that the initial Apl2 cDNA clone (G12) was shorter than the size of its corresponding mRNA. The cDNA libraries were rescreened and several longer clones were isolated, among them, two clones, A6 and B6, together with the original clone (G12) were further characterized (Figure 4). The longest open reading frame from the 2,492 nucleotides encodes a putative protein of 476 amino acids. Distinctive features of the putative protein include the C-terminus 100 amino acids that are very acidic, and an N-terminal 43 amino acid region dominated by 15 proline residues. Following the proline-rich region are typical leucine repeats (Landschulz, et al., Science 240:1759-1764 (1988); Vinson, et al., Science 246:911-916 (1989)), flanked by stretches of basic amino acids, suggesting a potential DNA-binding domain. These features are indicative of several different classes of eukaryotic transcription factors. In addition, a stretch of amino acids (LXSXE-----DDE) (SEQ ID NO: 1) at position 389-411 resembles the sequences of T-antigen which are responsible for binding to RB protein (DeCaprio, et al., 1988, supra.). Furthermore, there are two potential phosphorylation sites for Cdk kinase (Shenoy, et al., Cell 57:763-774 (1989)) at amino acids 159-161 (KSP) and 346-349 (SPGK) (SEQ ID NO: 2), which could modulate the function of this protein.

Apl2 binds only the hypophosphorylated form of RB at regions similar to those required for binding of SV40 T-antigen. To analyze the RB-binding properties of Apl2, the original clone (G12) was expressed as a Gst-fusion protein (P3) and purified by glutathione agarose chromatography. This fusion protein was used to test the binding of the Apl2 protein to full-length RB prepared from a cellular lysate of Molt4 cells, that expresses both hyper- and hypophosphorylated forms of the RB protein. Two additional controls were included in this experiment: one was a Gst-T-antigen fusion protein as a positive control and the other

was Gst alone as negative control. As shown in Figure 5A, the P3 protein binds only to the hypophosphorylated form and the binding affinity is very similar to that of T. Gst alone binds no detectable RB protein. To define which domain of RB is binding to Apl2, a panel of RB mutants expressed in the bacterial pET-T7 expression system (Studier et al., Meth. Enzymol. 185:60-89 (1990)) were mixed with the P3 beads or in parallel, with Gst-T beads. The amount of wild type or mutated RB proteins bound to the beads was determined by Western blot analysis using a monoclonal anti-RB antibody (mAb245). As shown in Fig 5C and 5D, the mutated RB defective in binding to T also failed to bind to Apl2. These results indicate that both Apl2 and T bind to the unphosphorylated form of RB in similar regions, showing that the Apl2-RB association is biologically significant.

~~RESULTS~~ The C-terminal region of Apl2 is required for binding to RB. Since the initial P3 fusion protein which contains 114 amino acids of Apl2 binds to RB, additional experiments were designed to map the region of Apl2 required for binding to RB. Four Gst-Apl2 fusion proteins with different N-terminal or C-terminal deletions were constructed, XH9 contains the entire coding sequence of the Apl2 cDNA and SH5 (from Sma I to Hind III) contains the C-terminal 314 amino acids. XX4 and SX4 are derived from XH9 and SH5, respectively, and contain a deletion of 21 amino acids at the C-terminus. The bacterially expressed RB protein (pETRbc) was mixed with these Gst-Apl2 derivatives and analyzed by Western blotting, as described above. Xh9, SH5 and P3 bind to RB with similar affinity, suggesting that the N-terminal sequence of Apl2 contributes little to RB-binding. However, XX4 and SX4, that both have 21 amino acids deleted from the C-terminus but contain the (LXSXE---DDE) sequence (DeCaprio, et al., 1988, supra.; Phelps, et al., J. Virol. 66:2418-2427 (1992)), failed to bind RB (Figure 6). Together, these results indicate that the C-

terminal region of Ap12 is required for binding to RB and the (LXSXE---DDE) sequence alone is not sufficient for binding, suggesting that the mode of RB-Ap12 interaction may be different from that of RB-T or RB-E1A interaction.

5 **Ap12 binds specifically to the E2F recognition sequence.** Since it has been shown that RB forms a complex with the transcription factor E2F (Bagchi, et al., Cell 65:1063-1072 (1991); Bandara, et al., Nature 352:249-251 (1991); Chellappan, et al., Cell 65:1053-1061 (1991)), and
10 Ap12 has a potential DNA-binding domain, experiments were performed to determine whether Ap12 could interact with an E2F binding site. The bacterially expressed Gst-Ap12 (SH5) fusion protein was used in the DNA mobility shift assay of a DNA fragment containing two E2F recognition sites using
15 previously described conditions (Yee, et al., Mol. Cell Biol. 9:578-585 (1989)). As shown in Figure 7A, SH5 binds that probe specifically since the complex is effectively
20 competed with the unlabeled DNA fragment containing the wild-type E2F cognate sequence but not by a mutated sequence that differs from the wild type by only two nucleotides (Yee, et al., supra.). As a positive control, partially purified E2F protein from HeLa cells specifically binds to the DNA probe as expected.

To determine if RB is able to interact with the
25 Ap12-DNA sequence specific complex, purified p 56-RB protein was included in the DNA mobility shift assay. The experiments were performed in two ways, either SH5 was mixed with RB then added to the E2F probe (Fig 7B, lane 3) or the fusion protein was bound to the E2F probe first
30 followed by addition of RB (Figure 7B, lane 6). In either case, the Ap12-DNA complex was super-shifted to more slowly migrating positions by adding RB, indicating that RB has the ability to interact with the specific Ap12-DNA complex. These results show that the Ap12 protein has a DNA-binding

as well as a RB-binding activity similar to that shown for E2F.

To determine whether the region containing the leucine repeats is required for DNA binding, three Gst-Ap12 fusion proteins, P3, SH5 and XH9 were chosen for DNA mobility shift assays. As shown in Figure 7C, SH5 and XH9 which contain the putative leucine zipper and stretches of basic amino acid residues (bZIP) (Vinson et al., supra.) bound to the E2F recognition sequence whereas the C-terminal region of Ap12 (P3) did not. In addition, some other controls, Ap9, Ap15 and Gst alone, also tested negative. This result demonstrates that a region containing the putative bZIP motif is necessary for the Ap12-DNA specific interaction.

The C-terminus of AP12 can function as a transactivation domain. Highly acidic, amphipathic alpha-helical regions commonly serve as activation domains in eukaryotic transcription factors (for review see Mitchell and Tjian, supra.). The C-terminal region of AP12 also displayed these characteristics, suggesting that it may function in an analogous manner. To test this, AP12 sequences encoding either amino acids 22-476 or the C-terminal 114 amino acids (362-476) were fused to those for the DNA binding domain of the yeast GAL4 protein (amino acids 1-147) (Keegan, et al., Science 231:699-704 (1986)) present on a yeast expression vector. While this GAL4 fragment can bind specifically to its recognition site (UAS_c) (Keegan, et al., supra.), it lacks an activation domain. Therefore, the chimeric protein relies on the fused segment to provide activation functions in order to direct transcription from a UAS_c containing promoter. Several such fusions involving mammalian activators have been shown to be functional in yeast, including p53 (Fields and Jang, Science 249:1046-1051 (1990)). As shown in Figure 8, following transformation of yeast strain

harboring the *E. coli lacZ* gene under UAS_c control, both GAL4-AP12 fusions were able to activate transcription of the reporter as evidenced by β -galactosidase activity wher as the GAL4-RB control was not. This result indicates
5 that AP12 does contain an activation domain, and that the C-terminal 114 amino acids are sufficient for this function.

Expression of Ap12 in CV1 cells transactivates a promoter with E2F recognition sequences. To determine
10 whether Ap12 can activate transcription in an E2F binding site-dependent manner, two plasmids, CMV-Ap12-Stu and CMV-Ap12-RH, were constructed to express the Ap12 in mammalian cells under the control of a cytomegalovirus(CMV)-IE promoter (Neill, et al., J. Virol. 65:5364-5373 (1991))
15 (Figure 9A). Two reporter plasmids, pE2FA₁₀CAT with two E2F sites upstream of the CAT reporter gene, and pA₁₀CAT ~~also~~ containing no E2F binding sites (Yee, et al., supra.), were ~~also~~ used for this assay. Figure 9B showed that the expression ~~trans~~ of either CMV-Ap12-Stu or CMV-Ap12-RH significantly
20 enhanced CAT activity when pE2FA₁₀CAT, but not pA₁₀CAT, was cotransfected. Expression of CMV-E4 has no apparent effect when compared with the control cells which were only transfected with the reporter plasmid. These data suggested that Ap12 encodes a functional transcription
25 factor which activates promoters with E2F recognition sequences.

Isolation of cellular genes encoding Rb-associated proteins. Two cDNA libraries were constructed from poly A⁺ RNA isolated from HeLa cells and Saos2 cells by
30 previously described methods (Sambrook et al., supra.). The double stranded cDNAs were size fractionated by using Sepharose Cl-4B chromatography and were ligated to λ gt11 arms. The size of the in vitro packaged libraries was 2.0×10^7 recombinants for HeLa cells and 1.5×10^7 for Saos2
35 cells with the average size of inserts being 1.6 kb. The

cdNA libraries were plated on one hundred 150mm dishes at $1-2 \times 10^4$ recombinants per dish and incubated at 42°C until plaques just became visible (3.5 hours), and then transferred to the nitrocellulose filters saturated with IPTG (10 mM) for overnight at 37°C. The filters were denatured and renatured in 6M guanidine HCl and incubated with the RB-sandwich probe in binding buffer (25 mM Hepes, pH 7.5, 50 mM NaCl, 5mM MgCl₂, 5 mM DTT, 0.1% NP-40, 5% milk, 1 mg/ml BSA) for 4 hours at 4°C. The RB-sandwich was prepared by mixing 1 µg of purified bacterially expressed p56-RB (Huang et al., 1991, supra.), 100 µl of preabsorbed polyclonal anti-RB antibody (anti-RB 0.47, 1:100 dilution) and 1 µl of alkaline-phosphatase conjugated secondary antibody (1:1000 dilution) per ml of binding buffer, incubated at 4°C for 2 hours. The RB-minus control sandwich was prepared by mixing the RB antibody and the secondary antibody and used as a control to eliminate the clones cross-reacted with the anti-RB antibody. The bound filters were then washed in TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween-20) 5 times, 3 minutes each and color developed in BCIP/NBP (Promega, WI). Positive clones from the initial screening were picked and subjected to second and third rounds of screening. The clones that consistently showed positive signals with the RB-sandwich but not with the RB-minus sandwich were then selected for fourth and fifth rounds of screening by plating at low density mixed with control phages to ensure homogenous isolates obtained which gave strong positive signals over the background.

Plasmid construction and fusion protein expression. The cdNA inserts of RbAps clones were subcloned into the pGEM1 for sequencing analysis. To express RbAp fusion proteins in vitro, the cdNA inserts were reconstructed in-frame into the pFLAG fusion protein expression system (IBI). The expression of the FLAG fusion proteins were induced by 0.2 mM of IPTG and the bacterial

lysates were prepared by two rounds of freeze-and-thaw followed by sonication in lysis buffer B (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM DTT, 0.2% NP-40, 1 mM PMSF, 1 μ g/ml Leupeptin, 5 μ g/ml Aprotinin, 1 μ g/ml Antipain) and
5 were clarified by centrifugation. To express the RB protein in vitro, the p56 version of the RB cDNA fragment (aa 377-928) was subcloned into a plasmid expressing glutathione S-transferase (GST) fusion protein pGEX-2T (Smith and Johnson, supra.) and the bacterially expressed
10 GST-RB fusion was prepared and purified using GST agarose beads.

In vitro binding assay. Bacterial lysates (100 μ l) containing about 0.5 μ g of the FLAG-RbAps were mixed with 20 μ l of the GST-RB beads or GST beads carrying 1-2 μ g
15 of the fusion protein in 400 μ l lysis buffer B at 4°C for 60 minutes. The bound beads were subsequently washed 5 times in 1 ml PBS/0.2%NP-40 and the protein complex was boiled in SDS loading buffer. The bound FLAG fusion proteins were then analyzed by SDS polyacrylamide gel
20 electrophoresis, immunoblotted and probed with an anti-FLAG monoclonal antibody (IBI).

Construction of mutated RB proteins expressed in the bacterial pET-T7 system. In addition to pETRbc, pETM6 and pETM9 (Huang et al., 1991, supra.), pETB2, pETSsp and
25 pETM8 were constructed by cloning AhaII-BamHI fragments from pB2, pSsp and pM8 (Huang et al., 1990, supra.) into the corresponding pET expression vector. The bacterial lysates were prepared as described in previous section.

Construction of GST-RbAp12 fusion proteins. The
30 DNA fragments derived from RbAp12 clones were subcloned into the GST fusion plasmids. GST-P3 was constructed by cloning the Eco RI-Sph I fragment from the original C-terminal 1.3kb cDNA (G12) into pGEPK, a derivative from pGEX-2T (Smith and Johnson, supra.). GST-SH5 contains the

SmaI-HindIII fragment from clone B6 and GST-XH9 contains the EcoRI-HindIII fragment of clone A6 that contains the entire coding sequence. GST-SX4 and GST-XX4 are derived from GST-SH5 and GST-XH9, respectively, but the C-terminal 5 XhoI-HindIII fragment is deleted.

RNA Blot Analysis. Total RNA extracted by the guanidine isothiocyanate-CsCl method (Sambrook et al., supra.) was denatured in 50% formamide, 2.2M formaldehyde, 20 mM Na borate (pH 8.3) and analyzed by 1.0% agarose gel electrophoresis. The RNA was then transferred to Hybond paper (Amersham) and the blot was immobilized by UV crosslinking. Prehybridization and hybridization were carried out in 50% formamide, 5x SSPE, 5x Denhardt's, 1% SDS and 100 µg/ml salmon sperm DNA and hybridization was performed in presence of ³²P-labeled 1.3 kb RbAp12 insert DNA at 45°C for 18 hours. The initial washing was carried out in 2x SSC, 0.1% SDS at room temperature and the final washing was in 0.1x SSC, 0.1% SDS at 65°C for 45 minutes.

DNA gel mobility shift assay. The insert from plasmid containing two E2F recognition sequences (TTTCGCGC--GCGCGAAA) (SEQ ID NO: 3) was used as a probe for the gel mobility shift assay and also served as a competitor. A plasmid containing a mutated E2F site (TTTAGCGC---GCGCTAAA) (SEQ ID NO: 4) (Huang et al., DNA and Cell Biol. 11:539-548 (1993)), which does not bind to E2F, was also used as a competitor. The assay was performed as described previously (Yee et al., supra.). The diluted GST-Ap12 bacterial lysates (20ng for SH5 and XH9 fusion proteins, 200ng for P3, Gst, GstAp9 and GstAp15) were incubated with 1x binding buffer (20 mM Hepes, pH 7.6, 1 mM MgCl₂, 0.1 mM EGTA, 40 mM KCl, 10% glycerol), 0.1% NP40, 1mg/ml salmon sperm DNA at room temperature for 15 minutes and the ³²P-end-labeled (Klenow fill-in) probe was added for another 30 minutes. The protein-DNA complexes were analyzed by 4% acrylamide gel electrophoresis in 0.25x TBE buffer at 4°C.

Yeast Expression Vector and Strain. The expression plasmid used in yeast was based on the pAS1 vector. Briefly, the plasmid contains the ADH1 promoter driving expression of the GAL4 DNA-binding domain followed by a downstream polylinker. The vector also carries the 2 μ origin and TRP1 gene for maintenance and selection in yeast. pAS/G12 was constructed by subcloning the EcoRI fragment isolated from G12 into the unique EcoRI site in pAS1. Similarly, pAS/12B6 was built using the EcoRI fragment from p12B6 and subcloning into the pAS1 EcoRI site. pASRb2 will be described elsewhere. The *Saccharomyces cerevisiae* strain used was Y153 (MATa, trp1-901, leu2-3, -112, ade2-101, ura3-52::URA3 (GAL1-lacZ), MEL (GAL1-lacZ)).

Yeast Transformation and β -galactosidase Assay. Yeast transformation was carried out using the LiOAc method as described previously (Schiestl and Gietz, Curr. Genet. 16:339-346 (1989)). After transformation, cells were plated on synthetic dropout media lacking tryptophan to select for the presence of the plasmid. Following 2-3 days growth at 30°C, single colonies from each transformation were streaked onto another selective plate and allowed to grow an additional 24 hours. The colony color β -galactosidase activity assay was then performed as described (Breen and Nasmyth, Quant. Biol. 50:643-650 (1985)) except the nitrocellulose filters were submerged in liquid nitrogen for about 30s-60s to permeabilize the cells, then thawed at room temperature before overlaying on Whatman filters saturated with LacZ-X-Gal solution (Breen and Nasmyth, supra.). The color developed in about 20 minutes in the case of the AP12 clones. No color change was observed with the pAS/Rb2 clone even after overnight exposure.

Transient Transfection Assay. The transfections were carried out with CV1 cells by conventional calcium

phosphate precipitation method. The plasmid pCMVAp12Stu was constructed by cloning the StuI fragment from clone A6 into the SmaI site of pCMV and plasmid pCMVAp12RH contains the EcoRI-HindIII fragment of clone B6. The plasmid pCMVE4
5 was used as a control. The CMV constructs were cotransfected with plasmids pE2FA₁₀CAT (containing two E2F binding sites) and pA₁₀CAT (containing no E2F binding sites) with the same number of cells (5×10^6) and the CAT activities were measured after 48 hours as described previously
10 (Gorman et al., Mol. Cell Biol. 2:1044-1051 (1982)).

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention.
15 Accordingly, the invention is limited only by the claims which follow.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM
- 5 (ii) TITLE OF INVENTION: CELLULAR GENES ENCODING
RETINOBLASTOMA-ASSOCIATED PROTEINS
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
- 10 (A) ADDRESSEE: CAMPBELL AND FLORES
(B) STREET: 4370 LA JOLLA VILLAGE DRIVE
(C) CITY: SAN DIEGO
(D) STATE: CALIFORNIA
(E) COUNTRY: USA
(F) ZIP: 92122
- 15 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- 20 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
(B) FILING DATE: 19-NOV-1993
(C) CLASSIFICATION:
- 25 (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: CAMPBELL, CATHRYN
(B) REGISTRATION NUMBER: 31,815
(C) REFERENCE/DOCKET NUMBER: FP-CJ 9790
- (ix) TELECOMMUNICATION INFORMATION:
- 30 (A) TELEPHONE: 619-535-9001
(B) TELEFAX: 619-535-8949

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- 35 (A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: internal
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- 40 Leu Xaa Ser Xaa Glu Asp Asp Glu
1 5

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- 45 (A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

34

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser Pro Gly Lys
1

5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTTCGCGCGC GCGAAA

16

15 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTAGCGCGC GCTAAA

16

(2) INFORMATION FOR SEQ ID NO:5:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCCTTGACC TTGCTGGGAA TGCTCGGTCA GACAAGGGCA GCATGTCTGA AGACTGTGGG 60

CCAGGAACCT CCGGGGAGCT GGGCGGCTGA GGCGATCAAA ATTGAGCCAG AGGATCTGGA 120

CATCATTCAG GTCACCGTCC CAGACCCCTC GCCAACCTCT GAGGAAATGA CAGACTCG 178

35 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 151 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTTTTTACTT ATTTAAAAAG GCCTTGGTGG CAGGAATATA GTGTAAAAAT CATTGGAAAA	60
ACTAAAAGGC ATCGATACAT ATCCGAATAT ACATTTTGTA CATAAATTAC ATTCCTTTA	120
GTCTTTCTGA GTGAGGTCCT GATTCAGTAC T	151

(2) INFORMATION FOR SEQ ID NO:7:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 255 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TTTACGACAG AGCACTATTG CCAAGCGTTC AAATGCAGCA CCATTAAGTA ACACAAAAAA	60
AGCATCTGGG AAGACTGTAT CTA CTGCTAA AGCAGGAGTG AAACAACCAG AAAGGAGTCA	120
GGTTAAAGAA GAAGTATGTA TG TCACTGAA ACCTGAGTAC CATAAGGAGA ATAGAAGGTG	180
20 CAGCCGAAAT AGCGGACAAA TTGAAGTGGA TACCTGAAGT ATCAGTGTCT TCAAGTCATT	240
CTTCAGTGTC ATCTT	255

(2) INFORMATION FOR SEQ ID NO:8:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

30 GAATTCAACT GTAGCTTGGT TTTCCAAAGT ATCTGGATCT AGTATTTTCTAG TCTTTTTGTGTC	60
TTCTTCAGCA CAACATTTTA CACAGACATA TTCTTTGTCT TCCTCGCCCA TCTGCTGTGC	120
TTGAGAAAGA CTTAACCCTAA CACAATCACC ATGAAACCAG TCATCACATC TCCACAGCCA	180
ACCATAACTG TTGCATGTGT TTTTGCAAAC CACACTGTTG CTGGAGTCAC ATATATTCGT	240
TCAAT	245

35 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

36

- (A) LENGTH: 688 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAATTCAGTG	GAGCACCAGT	AGAAGGTGCA	GGAGAAGAGG	CATTGACTCC	ATCAGTTCCT	60
ATAAATAAAG	GTCCCAAACC	TAAGAGGGAG	AAGAAGGAGC	CTGGTACCAG	AGTGAGAAAA	120
ACACCTACAT	CATCTGGTAA	ACCTAGTGCA	AAGAAAGTGA	AGAAACGGAA	TCCTTGGTCA	180
10 GATGATGAAT	CCAAGTCAGA	AAGTGATTTG	GAAGAAACAG	AACCTGTGGT	TATTCCAAGA	240
GATTCTTTGC	TTAGGAGAGC	AGCAGCCGAA	AGACCTAAAT	ACACATTTAA	TTTCTCAGAA	300
GAAGAGGATG	ATGATGCTGA	TGATGATGAT	GATGACAATA	ATGATTTAGA	GGAATTGAAA	360
GTTAAAGCAT	CTCCCAT AAC	AAATGATGGG	GAAGATGAAT	TTGTTCTTTC	AGATGGGTTA	420
GATAAAGATG	AATATACATT	TTCACCAGGC	AAATCAAAAG	CCTCACCAGA	AAAATCTTTG	480
15 CATGACAAAA	AAAGTCAGGA	TTTTTGAAAT	CTCTTCTCAT	TTCCTTCATA	TTCTCAGAAG	540
TCAGAAGATG	ATTCAGCTAA	ATTTGACAGT	AATGAAGAAG	ATTCTGCTTC	TGTTTTTTTCA	600
CCATCATTTG	GTCTGAAACA	GACAGATAAA	GTTCCAAGTA	AAACGGTAGC	TGCTAAAAAG	660
GGAAAACCGT	CTTCAGATAC	AGTCCCTA				688

(2) INFORMATION FOR SEQ ID NO:10:

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 348 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCAATGTTTA	ATTAAGTGGG	GAAAGAGCAC	AAACATTTTT	CAACAAATAC	TTGTGTTGTC	60
CTTTTGCTTT	CTCTGTCTCA	GACCTTTTGT	ACATCTGGCT	TATTTTAATG	TGATGATGTA	120
ATTGACCGTT	TTTTATTATT	GTGGTAGGCC	TTTAAACATT	TTGTTCTTAC	ACATACAGTT	180
30 TTATGCTCTT	TTTTACTCAT	TGAAATGTCA	CGTACTGTCT	GATTGGCTTG	TAGAATTGGT	240
TATAGACTGC	CGTGCATTAG	CACAGATTTT	AATTGTCATG	GTTACAAACT	ACAGACCTGC	300
TTTTTGAAAT	GAAATTTAAA	CATTAAAAAT	GGAAGTGTGA	AAAAAAA		348

(2) INFORMATION FOR SEQ ID NO:11:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1800 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

	GAATTCGCGG CCAAGAAGCC TAATGAGAAA AACAAACCAC TTGATAATAA GGGAGAAAAA	60
	AGAAAAAGAA AAACTGAAGA AAAAGGCGTA GATAAAGATT TTGAGTCTTC TTCAATGAAA	120
	ATCTCGAAAC TAGAAGTGAC TGAAATAGTG AAACCATCAC CAAAGCGCAA AATGGAACCT	180
	GATACTGAAA AAATGGATAG GACCCCTGAA AAGGACAAAA TTTCTTTAAG TGCGCCAGCC	240
10	AAAAAATCA AACTCAACAG AGAAACTGGG AAGAAAATTG GAAGTACAGA AAATATATCA	300
	AACACAAAAG AACCCCTCTGA AAAATTGGAG TCAACATCTA GCAAAGTTAA ACAAGAAAAA	360
	GTCAAAGGAA AGGTCAGACG AAAAGTGACT GGAAGTGAAG GATCCAGCTC AACTCTGGTG	420
	GATTACACCA GTACGAGCTC AACTGGAGGC AGTCCTGTGC GGAAATCTGA AGAAAAACA	480
	GATACAAAGC GAACTGTGAT TAAAAGCATG GAAGAATATA ATAATGACAA TACCGCGCCA	540
15	CGTGAAGATG TTATCATTAT GATTCAAGTT CCTCAATCCA AATGGGATAA AGATGACTTT	600
	GAATCTGAAG AAGAAGATGT TAAATCCACA CAGCCTATAT CAAGTGTAGG AAAACCTGCT	660
	AGTGTTATAA AAAATGTTAG TACAAAGCCA TCAAATATAG TCAAGTATCC TGAGAAAGAA	720
	AGTGAGCCAT CCGAGAAAAT TCAGAAATTC ACCAAGGACG TGAGCCATGA AATCATACAA	780
	CATGAGGTTA AAAGTTCAAA AAACCTCTGCA TCTAGTGAAA AAGGGAAAAC CAAAGATCGA	840
20	GATTATTCAG TGTGGAAGAA GGAGAACCCT GAAAAGAGGA AGAACAGCAC TCAGCCAGAG	900
	AAAGAGAGTA ATTTGGACCG TCTGAATGAA CAAGGAAATT TTAAAAGTCT GTCTCAATCT	960
	TCCAAAGAGG CTAGAACGTC AGATAAACAT GATTCCACTC GTGCTTCCTC AAATAAAGAC	1020
	TTCACTCCCA ATAGAGACAA AAAAAGTAC TATGACACCA GAGAGTATTC AAGTTCCAAA	1080
	CGTAGAGATG AAAAGAATGA ATTAACAAGA CGAAAAGACT CTCCTTCTCG GAATAAAGAT	1140
25	TCTGCATCTG GACAGAAAAA TAAACCAAGG GAAGAGAGAG ATTTGCCTAA AAAAGGAACA	1200
	GGAGATTCCA AAAAAAGTAA TTCTAGTCCC TCAAGAGACA GAAAACCTCA TGATCACAAA	1260
	GCCACTTATG ATACTAAACG GCCAAATGAA GAGACAAAAT CTGTAGATAA AAATCCTTGT	1320
	AAGGATCGTG AGAAGCATGT ATTAGAAGCA AGGAACAATA AAGAGTCAAG TGGCAATAAA	1380
	CTACTTTATA TACTTAACCC ACCAGAGACA CAGGTTGAAA AAGAGCAAAT TACTGGGCAA	1440
30	ATTGACAAGA GTACTGTCAA GCCTAAACCC CAGTTAAGTC ATTCCTCTAG ACTTTCCTCT	1500
	GACTTAACTA GAGAACTCA TGAAGCTGCT TTTGAACCAG ACTATAATGA AAGTGACAGT	1560
	GAAAGTAATG TTTCTGTAAA AGAAGAGGAA TCTTCAGGAA ACATTTCTAA GGACCTGAAA	1620
	GATAAAATAG TGGAGAAAGC AAAAGAGAGC CTGGACACAG CAGCAGTTGT CCAGGTGGGC	1680
	ATAAGCAGGA ATCAGAGCCA CAGCAGCCCC AGCGTCAGCC CCAGCAGAAG CCACAGTCCT	1740

TCTGGAAGCC AGACCCGAAG CCACAGTAGC AGTGCCAGCT CAGCAGAAAAG TCAGGACAGC 1800

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 4868 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

10	GAATTCCGGC CGGAATTAAT TCCGGGGATT TCCTGGGGAA TCAGGAAGAT ATCCATAATC	60
	TTCAACTGCG GGTAAAAGAG ACATCAAATG AGAATTTGAG ATTACTTCAT GTGATAGAGG	120
	ACCGTGACAG AAAAGTTGAA AGTTTGCTAA ATGAAATGAA AGAATTAGAC TCAAAACTCC	180
	ATTTACAGGA GGTACAATA ATGACCAAAA TTGAAGCATG CATAGAATTG GAAAAAATAG	240
	TTGGGGAACT TAAGAAAGAA AACTCAGATT TAAGTGAAAA ATTGGAATAT TTTTCTTGTG	300
15	ATCACCAGGA GTTACTCCAG AGAGTAGAAA CTTCTGAAGG CCTCAATTCT GATTTAGAAA	360
	TGCATGCAGA TAAATCATCA CGTGAAGATA TTGGAGATAA TGTGGCCAAG GTGAATGACA	420
	GCTGGAAGGA GAGATTTCTT GATGTGGAAA ATGAGCTGAG TAGGATCAGA TCGGAGAAAG	480
	CTAGCATTGA GCATGAAGCC CTCTACCTGG AGGCTGACTT AGAGGTAGTT CAAACAGAGA	540
	AGCTATGTTT AGAAAAAGAC AATGAAAATA AGCAGAAGGT TATTGTCTGC CTTGAAGAAG	600
20	AACTCTCAGT GGTCAACAAGT GAGAGAAACC AGCTTCGTGG AGAATTAGAT ACTATGTCAA	660
	AAAAAACCAC GGCCTGGAT CAGTTGTCTG AAAAAATGAA GGAGAAAACA CAAGAGCTTG	720
	AGTCTCATCA AAGTGAGTGT CTCCATTGCA TTCAGGTGGC AGAGGCAGAG GTGAAGGAAA	780
	AGACGGAACCT CCTTCAGACT TTGTCCTCTG ATGTGAGTGA GCTGTTAAAA GACAAACTC	840
	ATCTCCAGGA AAAGCTGCAG AGTTTGGAAG AGGACTCACA GGCCTGTCTT TTGACAAAAT	900
25	GTGAGCTGGA AAACCAAATT GCACAACTGA ATAAAGAGAA AGAATTGCTT GTCAAGGAAT	960
	CTGAAAGCCT GCAGGCCAGA CTGAGTGAAT CAGATTATGA AAAGCTGAAT GTCTCCAAGG	1020
	CCTTGAGAGC CGCACTGGTG GAGAAAGGTG AGTTCGCATT GAGGCTGAGC TCAACACAGG	1080
	AGGAAGTGCA TCAGCTGAGA AGAGGCATCG AGAAACTGAG AGTTCGCATT GAGGCCGATG	1140
	AAAAGAAGCA GCTGCACATC GCAGAGAAAC TGAAAGAACG CGAGCGGGAG AATGATTCAC	1200
30	TTAAGGTAAA AGTTGAGAAC CTTGAAAGGG AATTGCAGAT GTCAGAAGAA AACCAGGAGC	1260
	TAGTGATTCT TGATGCCGAG AATTCCAAAG CAGAAGTAGA GACTCTAAAA ACACAAATAG	1320
	AAGAGATGGC CAGAAGCCTG AAAGTTTTTG AATTAGACCT TGTACGTTA AGGTCTGAAA	1380
	AAGAAAATCT GACAAAACAA ATACAAGAAA AACAAGGTCA GTTGTGAGAA CTAGACAAGT	1440
	TACTCTCTTC ATTTAAAAGT CTGTTAGAAG AAAAGGAGCA AGCAGAGATA CAGATCAAAG	1500

	AAGAATCTAA	AACTGCAGTG	GAGATGCTTC	AGAATCAGTT	AAAGGAGCTA	AATGAGGCAG	1560
	TAGCAGCCTT	GTGTGGTGAC	CAAGAAATTA	TGAAGGCCAC	AGAACAGAGT	CTAGACCCAC	1620
	CAATAGAGGA	AGAGCATCAG	CTGAGAAATA	GCATTGAAAA	GCTGAGAGCC	CGCCTAGAAG	1680
	CTGATGAAAA	GAAGCAGCTC	TGTGTCTTAC	AACAACCTGAA	GGAAAGTGAG	CATCATGCAG	1740
5	ATTTACTTAA	GGGTAGAGTG	GAGAACCTTG	AAAGAGAGCT	AGAGATAGCC	AGGACAAACC	1800
	AAGAGCATGC	AGCTCTTGAG	GCAGAGAATT	CCAAAGGAGA	GGTAGAGACC	CTAAAAGCAA	1860
	AAATAGAAGG	GATGACCCAA	AGTCTGAGAG	GTCTGGAATT	AGATGTTGTT	ACTATAAGGT	1920
	CAGAAAAAGA	AAATCTGACA	AATGAATTAC	AAAAAGAGCA	AGAGCGAATA	TCTGAATTAG	1980
	AAATAATAAA	TTCATCATTT	GAAAATATTT	TGCAAGAAAA	AGAGCAAGAG	AAAGTACAGA	2040
10	TGAAAGAAAA	ATCAAGCACT	GCCATGGAGA	TGCTTCAAAC	ACAATTAAAA	GAGCTCAATG	2100
	AGAGAGTGGC	AGCCCTGCAT	AATGACCAAG	AAGCCTGTAA	GGCCAAAGAG	CAGAATCTTA	2160
	GTAGTCAAGT	AGAGTGCTCT	GAACTTGAGA	AGGCTCAGTT	GCTACAAGGC	CTTGATGAGG	2220
	CCAAAAATAA	TTATATTGTT	TTGCAATCTT	CAGTGAATGG	CCTCATTCAA	GAAGTAGAAG	2280
	ATGGCAAGCA	GAAACTGGAG	AAGAAGGATG	AAGAAATCAG	TAGACTGAAA	AATCAAATTC	2340
15	AAGACCAAGA	GCAGCTTGTC	TCTAACTGT	CCCAGGTGGA	AGGAGAGCAC	CAACTTTGGA	2400
	AGGAGCAAAA	CTTAGAACTG	AGAAATCTGA	CAGTGGGAATT	GGAGCAGAAG	ATCCAAGTGC	2460
	TACAATCCAA	AAATGCCTCT	TTGCAGGACA	CATTAGAAGT	GCTGCAGAGT	TCTTACAAGA	2520
	ATCTAGAGAA	TGAGCTTGAA	TTGACAAAAA	TGGACAAAAT	GTCCTTTGTT	GAAAAAGTAA	2580
	ACAAAATGAC	TGCAAAGGAA	ACTGAGCTGC	AGAGGGAAAT	GCATGAGATG	GCACAGAAAA	2640
20	CAGCAGAGCT	GCAAGAAGAA	CTCAGTGGAG	AGAAAAATAG	GCTAGCTGGA	GAGTTGCAGT	2700
	TACTGTTGGA	AGAAATAAAG	AGCAGCAAAG	ATCAATTGAA	GGAGCTCACA	CTAGAAAATA	2760
	GTGAATTGAA	GAAGAGCCTA	GATTGCATGC	ACAAAGACCA	GGTGGAAAAG	GAAGGGAAAG	2820
	TGAGAGAGGA	AATAGCTGAA	TATCAGCTAC	GGCTTCATGA	AGCTGAAAAG	AAACACCAGG	2880
	CTTTGCTTTT	GGACACAAAC	AAACAGTATG	AAGTAGAAAT	CCAGACATAC	CGAGAGAAAT	2940
25	TGACTTCTAA	AGAAGAATGT	CTCAGTTCAC	AGAAGCTGGA	GATAGACCTT	TTAAAGTCTA	3000
	GTAAAGAAGA	GCTCAATAAT	TCATTGAAAG	CTACTACTCA	GATTTTGGAA	GAATTGAAGA	3060
	AAACCAAGAT	GGACAATCTA	AAATATGTAA	ATCAGTTGAA	GAAGGAAAAT	GAACGTGCCC	3120
	AGGGGAAAAAT	GAAGTTGTTG	ATCAAATCCT	GTAAACAGCT	GGAAGAGGAA	AAGGAGATAC	3180
	TGCAGAAAGA	ACTCTCTCAA	CTTCAAGCTG	CACAGGAGAA	GCAGAAAACA	GGTACTGTTA	3240
30	TGGATACCAA	GGTCGATGAA	TTAACAACCTG	AGATCAAAGA	ACTGAAAGAA	ACTCTTGAAG	3300
	AAAAAACCBA	GGAGGCAGAT	GAATACTTGG	ATAAGTACTG	TTCCTTGCTT	ATAAGCCATG	3360
	AAAAGTTAGA	GAAAGCTAAA	GAGATGTTAG	AGACACAAGT	GGCCCATCTG	TGTTACAGC	3420
	AATCTAAACA	AGATTCCCGA	GGGTCTCCTT	TGCTAGGTCC	AGTTGTTCCA	GGACCATCTC	3480
	CAATCCCTTC	TGTTACTGAA	AAGAGGTTAT	CATCTGGCCA	AAATAAAGCT	TCAGGCAAGA	3540

40

```

GGCAAAGATC CAGTGAATA TGGGAGAATG GTGGAGGACC AACACCTGCT ACCCCAGAGA 3600
CCTTTTCTAA AAAAAGCAAG AAAGCAGTCA TGAGTGGTAT TCACCCTGCA GAAGACACGG 3660
AAGGTACTGA GTTTGAGCCA GAGGGACTTC CAGAAGTTGT AAAGAAAGGG TTTGCTGACA 3720
TCCCGACAGG AAAGACTAGC CCATATATCC TGCGAAGAAC AACCATGGCA ACTGGGAGCA 3780
5  GGCCCGGCCT GGCTGCACAC AAGTTACCCC TATCCCCACT GACTGTCCCC AAACAAAATC 3840
TTGCAGAGTC CTCCAAACCA ACAGCTGGTG GCAGCAGATC ACAAAGGTG AAAGTTGCTC 3900
AGCGGAGCCC AGTAGATTCA GGCACCATCC TCCGAGAACC CACCACGAAA TCCGTCCCAG 3960
TCAATAATCT TCCTGAGAGA AGTCCGACTG ACAGCCCCAG AGAGGGCCTG AGGGTCAAGC 4020
GCCGGCGACT TGTCCCCAGC CCCAAAGCTG GACTGGAGTC CAAGGGCAGT GAGAACTGTA 4080
10 AGGTCCAGTG AAGGCACTTT GTGTGTCAGT ACCCCTGGGA GGTGCCAGTC ATTGAATAGA 4140
TAAGGCTGTG CCTACAGGAC TTCTCTTTAG TCAGGGCATG CTTTATTAGT GAGGAGAAAA 4200
CAATTCCTTA GAAGTCTTAA ATATATTGTA CTCTTTAGAT CTCCCATGTG TAGGTATTGA 4260
AAAAGTTTGG AAGCACTGAT CACCTGTTAG CATTGCAATT CCTCTACTGC AATGTAAATA 4320
GTATAAAGCT ATGTATATAA AGCTTTTGG TAATATGTTA CAATTAAAAT GACAAGCACT 4380
15 ATATCACAAT CTCTGTTTGT ATGTGGGTTT TACACTAAAA AAATGCAAAA CACATTTTAT 4440
TCTTCTAATT AACAGCTCCT AGGAAAATGT AGACTTTTGC TTTATGATAT TCTATCTGTA 4500
GTATGAGGCA TGAATAGTT TTGTATCGGG AATTCTCAG AGCTGAGTAA AATGAAGGAA 4560
AAGCATGTTA TGTGTTTTTA AGGAAAATGT GCACACATAT ACATGTAGGA GTGTTTATCT 4620
TTCTCTTACA ATCTGTTTTA GACATCTTTG CTTATGAAAC CTGTACATAT GTGTGTGTGG 4680
20 GTATGTGTTT ATTTCCAGTG AGGGCTGCAG GCTTCCTAGA GGTGTGCTAT ACCATGCGTC 4740
TGTCGTTGTG CTTTTTCTG TTTTATAGACC AATTTTAC AGTTCCTTGG TAAGCATTGT 4800
CGTATCTGGT GATGGATTAA CATATAGCCT TTGTTTTCTA ATAAAATAGT CGCCTTCGTA 4860
AAAAAAA 4868

```

(2) INFORMATION FOR SEQ ID NO:13:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2492 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- 30 (ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1428

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

35 CTT TGC AGG CAG CGG CGG CCG GGG GCG GAG CGG GAT CGA GCC CTC GCC 48
   Leu Cys Arg Gln Arg Arg Pro Gly Ala Glu Arg Asp Arg Ala Leu Ala
    1             5             10             15

```

41

	GAG	GCC	TGC	CGC	CAT	GGG	CCC	GCG	CCG	CCG	CCG	CCG	CCT	GTC	ACC	CGG	96
	Glu	Ala	Cys	Arg	His	Gly	Pro	Ala	Pro	Pro	Pro	Pro	Pro	Val	Thr	Arg	
				20					25					30			
5	GCC	GCG	CGG	GCC	GTG	AGC	GTC	ATG	GCC	TTG	GCC	GGG	GCC	CCT	GCG	GGC	144
	Ala	Ala	Arg	Ala	Val	Ser	Val	Met	Ala	Leu	Ala	Gly	Ala	Pro	Ala	Gly	
			35					40					45				
	GGC	CCA	TGC	GCG	CCG	GCG	CTG	GAG	GCC	CTG	CTC	GGG	GCC	GGC	GCG	CTG	192
	Gly	Pro	Cys	Ala	Pro	Ala	Leu	Glu	Ala	Leu	Leu	Gly	Ala	Gly	Ala	Leu	
			50				55					60					
10	CGG	CTG	CTC	GAC	TCC	TCG	CAG	ATC	GTC	ATC	ATC	TCC	GCC	GCG	CAG	GAC	240
	Arg	Leu	Leu	Asp	Ser	Ser	Gln	Ile	Val	Ile	Ile	Ser	Ala	Ala	Gln	Asp	
			65			70					75					80	
15	GCC	AGC	GCC	CCG	CCG	GCT	CCC	ACC	GGC	CCC	GCG	GCG	CCC	GCC	GCC	GGC	288
	Ala	Ser	Ala	Pro	Pro	Ala	Pro	Thr	Gly	Pro	Ala	Ala	Pro	Ala	Ala	Gly	
					85					90					95		
	CCC	TGC	GAC	CCT	GAC	CTG	CTG	CTC	TTC	GCC	ACA	CCG	CAG	GCG	CCC	CGG	336
	Pro	Cys	Asp	Pro	Asp	Leu	Leu	Leu	Phe	Ala	Thr	Pro	Gln	Ala	Pro	Arg	
				100					105					110			
20	CCC	ACA	CCC	AGT	GCG	CCG	CGG	CCC	GCG	CTC	GGC	CGC	CCG	CCG	GTG	AAG	384
	Pro	Thr	Pro	Ser	Ala	Pro	Arg	Pro	Ala	Leu	Gly	Arg	Pro	Pro	Val	Lys	
			115				120						125				
	CGG	AGG	CTG	GAC	CTG	GAA	ACT	GAC	CAT	CAG	TAC	CTG	GCC	GAG	AGC	AGT	432
	Arg	Arg	Leu	Asp	Leu	Glu	Thr	Asp	His	Gln	Tyr	Leu	Ala	Glu	Ser	Ser	
			130				135					140					
25	GGG	CCA	GCT	CGG	GGC	AGA	GGC	CGC	CAT	CCA	GGA	AAA	GGT	GTG	AAA	TCC	480
	Gly	Pro	Ala	Arg	Gly	Arg	Gly	Arg	His	Pro	Gly	Lys	Gly	Val	Lys	Ser	
					150						155					160	
30	CCG	GGG	GAG	AAG	TCA	CGC	TAT	GAG	ACC	TCA	CTG	AAT	CTG	ACC	ACC	AAG	528
	Pro	Gly	Glu	Lys	Ser	Arg	Tyr	Glu	Thr	Ser	Leu	Asn	Leu	Thr	Thr	Lys	
				165						170					175		
	CGC	TTC	CTG	GAG	CTG	CTG	AGC	CAC	TCG	GCT	GAC	GGT	GTC	GTC	GAC	CTG	576
	Arg	Phe	Leu	Glu	Leu	Leu	Ser	His	Ser	Ala	Asp	Gly	Val	Val	Asp	Leu	
				180					185					190			
35	AAC	TGG	GCT	GCC	GAG	GTG	CTG	AAG	GTG	CAG	AAG	CGG	CGC	ATC	TAT	GAC	624
	Asn	Trp	Ala	Ala	Glu	Val	Leu	Lys	Val	Gln	Lys	Arg	Arg	Ile	Tyr	Asp	
			195				200						205				
	ATC	ACC	AAC	GTC	CTT	GAG	GGC	ATC	CAG	CTC	ATT	GCC	AAG	AAG	TCC	AAG	672
	Ile	Thr	Asn	Val	Leu	Glu	Gly	Ile	Gln	Leu	Ile	Ala	Lys	Lys	Ser	Lys	
			210				215					220					
40	AAC	CAC	ATC	CAG	TGG	CTG	GGC	AGC	CAC	ACC	ACA	GTG	GGC	GTC	GGC	GGA	720
	Asn	His	Ile	Gln	Trp	Leu	Gly	Ser	His	Thr	Thr	Val	Gly	Val	Gly	Gly	
			225			230					235					240	
45	CGG	CTT	GAG	GGG	TTG	ACC	CAG	GAC	CTC	CGA	CAG	CTG	CAG	GAG	AGC	GAG	768
	Arg	Leu	Glu	Gly	Leu	Thr	Gln	Asp	Leu	Arg	Gln	Leu	Gln	Glu	Ser	Glu	
				245						250					255		
	CAG	CAG	CTG	GAC	CAC	CTG	ATG	AAT	ATC	TGT	ACT	ACG	CAG	CTG	CGC	CTG	816
	Gln	Gln	Leu	Asp	His	Leu	Met	Asn	Ile	Cys	Thr	Thr	Gln	Leu	Arg	Leu	
				260					265					270			
50	CTC	TCC	GAG	GAC	ACT	GAC	AGC	CAG	CGC	CTG	GCC	TAC	GTG	ACG	TGT	CAG	864
	Leu	Ser	Glu	Asp	Thr	Asp	Ser	Gln	Arg	Leu	Ala	Tyr	Val	Thr	Cys	Gln	
			275					280					285				

42

	GAC CTT CGT AGC ATT GCA GAC CCT GCA GAG CAG ATG GTT ATG GTG ATC Asp Leu Arg Ser Ile Ala Asp Pro Ala Glu Gln Met Val Met Val Ile 290 295 300	912
5	AAA GCC CCT CCT GAG ACC CAG CTC CAA GCC GTG GAC TCT TCG GAG AAC Lys Ala Pro Pro Glu Thr Gln Leu Gln Ala Val Asp Ser Ser Glu Asn 305 310 315 320	960
	TTT CAG ATC TCC CTT AAG AGC AAA CAA GGC CCG ATC GAT GTT TTC CTG Phe Gln Ile Ser Leu Lys Ser Lys Gln Gly Pro Ile Asp Val Phe Leu 325 330 335	1008
10	TGC CCT GAG GAG ACC GTA GGT GGG ATC AGC CCT GGG AAG ACC CCA TCC Cys Pro Glu Glu Thr Val Gly Gly Ile Ser Pro Gly Lys Thr Pro Ser 340 345 350	1056
15	CAG GAG GTC ACT TCT GAG GAG GAG AAC AGG GCC ACT GAC TCT GCC ACC Gln Glu Val Thr Ser Glu Glu Glu Asn Arg Ala Thr Asp Ser Ala Thr 355 360 365	1104
	ATA GTG TCA CCA CCA CCA TCA TCT CCC CCC TCA TCC CTC ACC ACA GAT Ile Val Ser Pro Pro Pro Ser Ser Pro Pro Ser Ser Leu Thr Thr Asp 370 375 380	1152
20	CCC AGC CAG TCT CTA CTC AGC CTG GAG CAA GAA CCG CTG TTG TCC CGG Pro Ser Gln Ser Leu Leu Ser Leu Glu Gln Glu Pro Leu Leu Ser Arg 385 390 395 400	1200
	ATG GGC AGC CTG CGG GCT CCC GTG GAC GAG GAC CGC CTG TCC CCG CTG Met Gly Ser Leu Arg Ala Pro Val Asp Glu Asp Arg Leu Ser Pro Leu 405 410 415	1248
25	GTG GCG GCC GAC TCG CTC CTG GAG CAT GTG CGG GAG GAC TTC TCC GGC Val Ala Ala Asp Ser Leu Leu Glu His Val Arg Glu Asp Phe Ser Gly 420 425 430	1296
30	CTC CTC CCT GAG GAG TTC ATC AGC CTT TCC CCA CCC CAC GAG GCC CTC Leu Leu Pro Glu Glu Phe Ile Ser Leu Ser Pro Pro His Glu Ala Leu 435 440 445	1344
	GAC TAC CAC TTC GGC CTC GAG GAG GGC GAG GGC ATC AGA GAC CTC TTC Asp Tyr His Phe Gly Leu Glu Glu Gly Glu Gly Ile Arg Asp Leu Phe 450 455 460	1392
35	GAC TGT GAC TTT GGG GAC CTC ACC CCC CTG GAT TTC TGACAGGGCT Asp Cys Asp Phe Gly Asp Leu Thr Pro Leu Asp Phe 465 470 475	1438
	TGGAGGGACC AGGGTTTCCA GAGATGCTCA CTTGTCTCT GCAGCCCTGG AGCCCCCTGT	1498
	CCCTGGCCGT CCTCCCAGCC TGTTTGGAATTAATTT ATACCCCTCT CCTCTGTCTC	1558
	CAGAAGCTTC TAGCTCTGGG GTCTGGCTAC CGCTAGGAGG CTGAGCAAGC CAGGAAGGGA	1618
40	AGGAGTCTGT GTGGTGTGTA TGTGCATGCA GCCTACACCC ACACGTGTGT ACCGGGGGTG	1678
	AATGTGTGTG AGCATGTGTG TGTGCATGTA CCGGGGAATG AAGGTGAACA TACACCTCTG	1738
	TGTGTGCACT GCAGACACGC CCCAGTGTGT CCACATGTGT GTGCATGAGT CCATGTGTGC	1798
	GCGTGGGGGG GCTCTAACTG CACTTTCGGC CCTTTTGCTC TGGGGGTCCC ACAAGGCCCA	1858
	GGGCAGTGCC TGCTCCAGCA ATCTGGTGCT CTGACCAGGC CAGGTGGGGA GGCTTTGGCT	1918
45	GGCTGGGCGT GTAGGACGGT GAGAGCACTT CTGTCTTAAA GGTTTTTTCT GATTGAAGCT	1978
	TTAATGGAGC GTTATTTATT TATCGAGGCC TCTTTGGTGA GCCTGGGGAA TCAGCAAAGG	2038

43

GGAGGAGGGG TGTGGGGTTG ATACCCCAAC TCCCTCTACC CTTGAGCAAG GGCAGGGGTC 2098
 CCTGAGCTGT TCTTCTGCCC CATACTGAAG GAACTGAGGC CTGGGTGATT TATTTATTGG 2158
 GAAAGTGAGG GAGGGAGACA GACTGACTGA CAGCCATGGG TGGTCAGATG GTGGGGTGGG 2218
 CCCTCTCCAG GGGGCCAGTT CAGGGCCCCA GCTGCCCCC AGGATGGATA TGAGATGGGA 2278
 5 GAGGTGAGTG GGGGACCTTC ACTGATGTGG GCAGGAGGGG TGGTGAAGGC CTCCCCCAGC 2338
 CCAGACCCTG TGGTCCCTCC TGCAGTGTCT GAAGCGCCTG CCTCCCCACT GCTCTGCCCC 2398
 ACCCTCCAAT CTGCACTTTG ATTTGCTTCC TAACAGCTCT GTTCCCTCCT GCTTTGGTTT 2458
 TAATAAATAT TTTGATGACG TTAATAAAAA AAAA 2492

(2) INFORMATION FOR SEQ ID NO:14:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 476 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Leu Cys Arg Gln Arg Arg Pro Gly Ala Glu Arg Asp Arg Ala Leu Ala
 1 5 10 15
 Glu Ala Cys Arg His Gly Pro Ala Pro Pro Pro Pro Val Thr Arg
 20 25 30
 20 Ala Ala Arg Ala Val Ser Val Met Ala Leu Ala Gly Ala Pro Ala Gly
 35 40 45
 Gly Pro Cys Ala Pro Ala Leu Glu Ala Leu Leu Gly Ala Gly Ala Leu
 50 55 60
 25 Arg Leu Leu Asp Ser Ser Gln Ile Val Ile Ile Ser Ala Ala Gln Asp
 65 70 75 80
 Ala Ser Ala Pro Pro Ala Pro Thr Gly Pro Ala Ala Pro Ala Ala Gly
 85 90 95
 Pro Cys Asp Pro Asp Leu Leu Leu Phe Ala Thr Pro Gln Ala Pro Arg
 100 105 110
 30 Pro Thr Pro Ser Ala Pro Arg Pro Ala Leu Gly Arg Pro Pro Val Lys
 115 120 125
 Arg Arg Leu Asp Leu Glu Thr Asp His Gln Tyr Leu Ala Glu Ser Ser
 130 135 140
 35 Gly Pro Ala Arg Gly Arg Gly Arg His Pro Gly Lys Gly Val Lys Ser
 145 150 155 160
 Pro Gly Glu Lys Ser Arg Tyr Glu Thr Ser Leu Asn Leu Thr Thr Lys
 165 170 175
 Arg Phe Leu Glu Leu Leu Ser His Ser Ala Asp Gly Val Val Asp Leu
 180 185 190
 40 Asn Trp Ala Ala Glu Val Leu Lys Val Gln Lys Arg Arg Ile Tyr Asp
 195 200 205

44

Ile Thr Asn Val Leu Glu Gly Ile Gln Leu Ile Ala Lys Lys Ser Lys
 210 215 220
 Asn His Ile Gln Trp Leu Gly Ser His Thr Thr Val Gly Val Gly Gly
 225 230 235 240
 5 Arg Leu Glu Gly Leu Thr Gln Asp Leu Arg Gln Leu Gln Glu Ser Glu
 245 250 255
 Gln Gln Leu Asp His Leu Met Asn Ile Cys Thr Thr Gln Leu Arg Leu
 260 265 270
 10 Leu Ser Glu Asp Thr Asp Ser Gln Arg Leu Ala Tyr Val Thr Cys Gln
 275 280 285
 Asp Leu Arg Ser Ile Ala Asp Pro Ala Glu Gln Met Val Met Val Ile
 290 295 300
 Lys Ala Pro Pro Glu Thr Gln Leu Gln Ala Val Asp Ser Ser Glu Asn
 305 310 315 320
 15 Phe Gln Ile Ser Leu Lys Ser Lys Gln Gly Pro Ile Asp Val Phe Leu
 325 330 335
 Cys Pro Glu Glu Thr Val Gly Gly Ile Ser Pro Gly Lys Thr Pro Ser
 340 345 350
 20 Gln Glu Val Thr Ser Glu Glu Glu Asn Arg Ala Thr Asp Ser Ala Thr
 355 360 365
 Ile Val Ser Pro Pro Pro Ser Ser Pro Pro Ser Ser Leu Thr Thr Asp
 370 375 380
 Pro Ser Gln Ser Leu Leu Ser Leu Glu Gln Glu Pro Leu Leu Ser Arg
 385 390 395 400
 25 Met Gly Ser Leu Arg Ala Pro Val Asp Glu Asp Arg Leu Ser Pro Leu
 405 410 415
 Val Ala Ala Asp Ser Leu Leu Glu His Val Arg Glu Asp Phe Ser Gly
 420 425 430
 30 Leu Leu Pro Glu Glu Phe Ile Ser Leu Ser Pro Pro His Glu Ala Leu
 435 440 445
 Asp Tyr His Phe Gly Leu Glu Glu Gly Glu Gly Ile Arg Asp Leu Phe
 450 455 460
 Asp Cys Asp Phe Gly Asp Leu Thr Pro Leu Asp Phe
 465 470 475

WE CLAIM:

1. An isolated nucleic acid molecule encoding a retinoblastoma-associated polypeptide.
2. The isolated nucleic acid molecule of claim 1, wherein the encoded retinoblastoma-associated polypeptide has transcriptional factor E2F biological activity.
3. The isolated nucleic acid molecule of claim 1, wherein the encoded retinoblastoma-associated polypeptide has RB-binding activity.
4. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule, a cDNA molecule or an RNA molecule.
5. An isolated nucleic acid molecule that hybridizes under stringent conditions to the isolated nucleic acid molecule of claim 1.
6. An isolated and purified polypeptide encoded by the nucleic acid molecule of claim 1.
7. An isolated and purified polypeptide encoded by the nucleic acid molecule of claim 2.
8. An isolated and purified polypeptide encoded by the nucleic acid molecule of claim 3.
9. A vector comprising the isolated nucleic acid molecule of claim 1.
10. A plasmid comprising the vector of claim 9.
11. A virus comprising the vector of claim 9.

12. A host cell comprising the vector of claim 9.

13. The host cell of claim 12, wherein the host cell is a bacterium, a yeast cell or a mammalian cell.

5 14. An antibody capable of specifically binding to a retinoblastoma-associated polypeptide present in the nucleus of the cell.

15. An immunologically reactive polypeptide fragment of the antibody of claim 14.

10 16. The antibody of claim 14, wherein the antibody is a monoclonal antibody.

17. The antibody of claim 14, wherein said antibody is labelled with a detectable marker.

15 18. A hybridoma cell line producing the antibody of claim 17.

19. A method for detecting a retinoblastoma-associated protein in a sample comprising: a. contacting the antibody of claim 14 with the sample under conditions permitting formation of an antibody-antigen complex; b.
20 detecting the presence of any complex so formed; c. the presence of complex indicating the presence of retinoblastoma-associated protein in the sample.

20. A method of recombinantly producing a retinoblastoma-associated protein which comprises growing
25 the host cell of claim 12 under suitable conditions permitting production of the protein and recovering and purifying the resulting protein so produced.

21. Th recombinantly produced protein of claim
- 20.

1 / 17

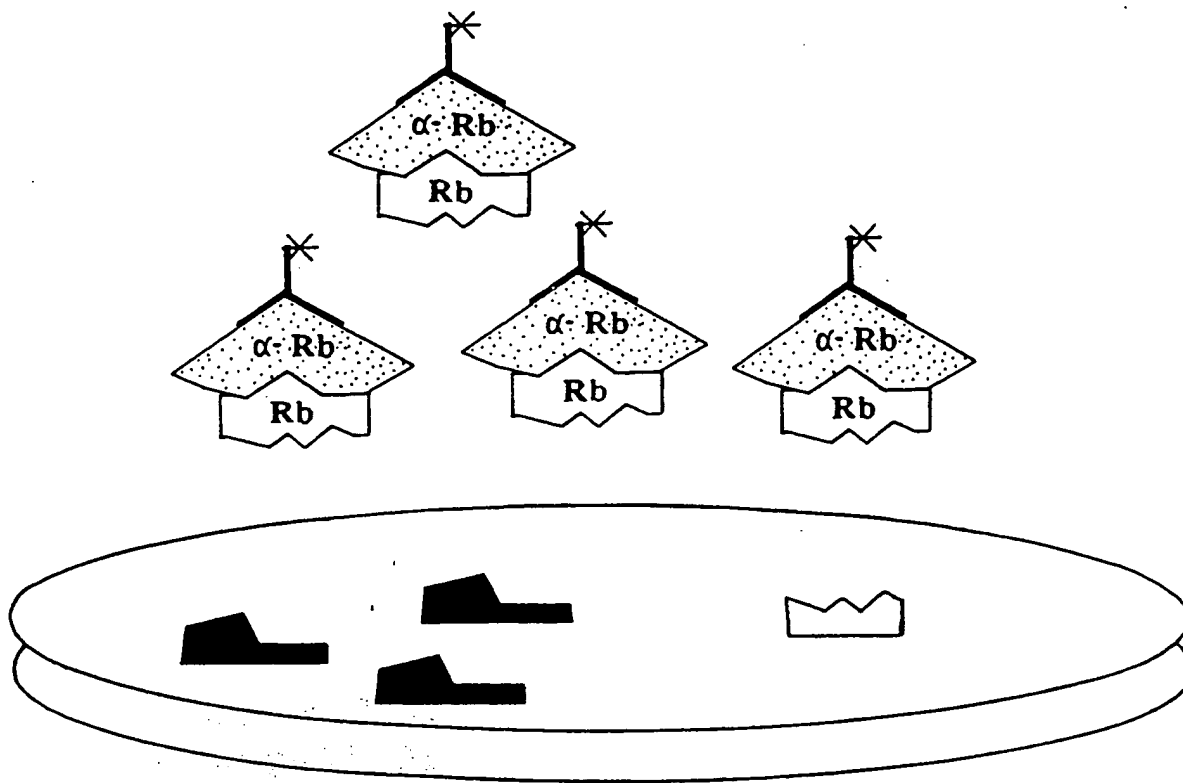


FIG. 1A

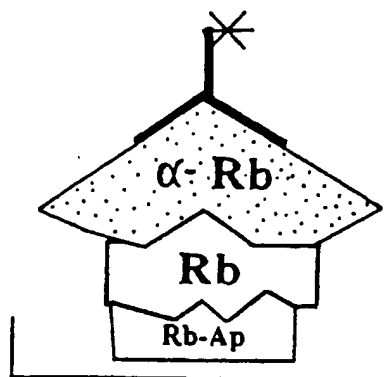


FIG. 1B

SUBSTITUTE SHEET

2 / 17

Rb-Ap

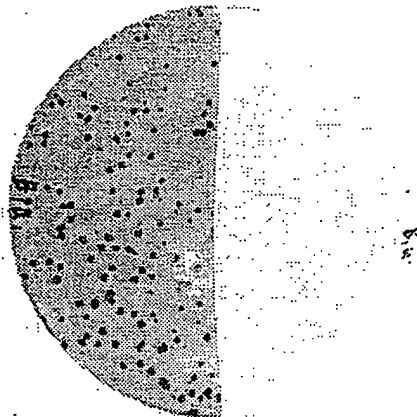


FIG. 1C

T antigen

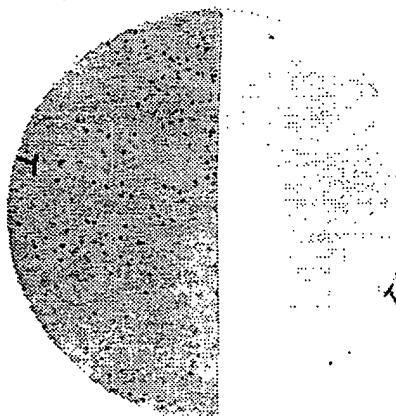


FIG. 1D

UBSTITUTE SHEET

3/17

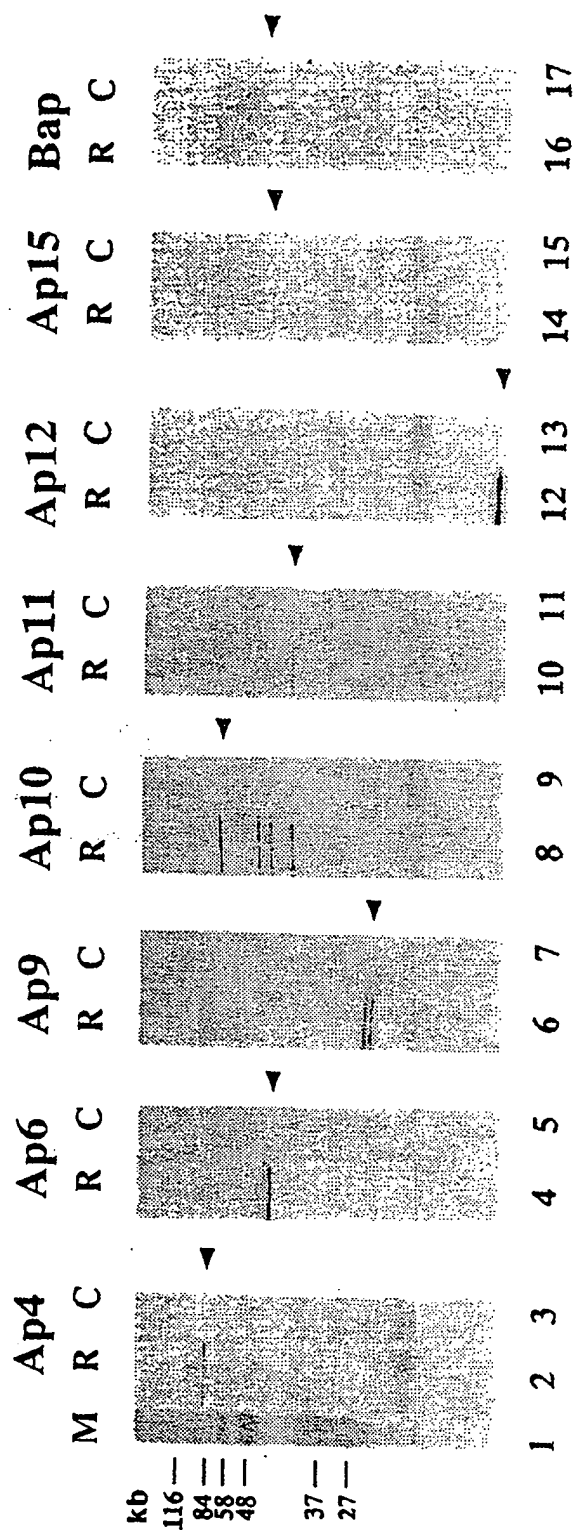


FIG. 2

4 / 17

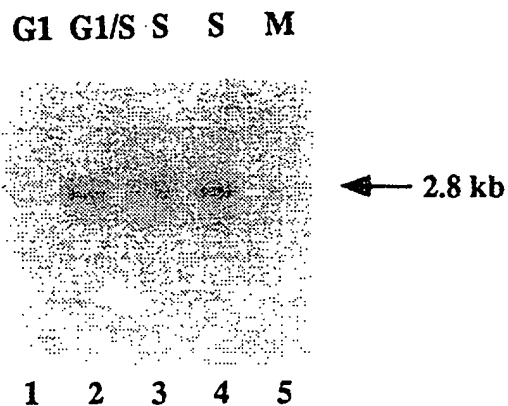


FIG. 3

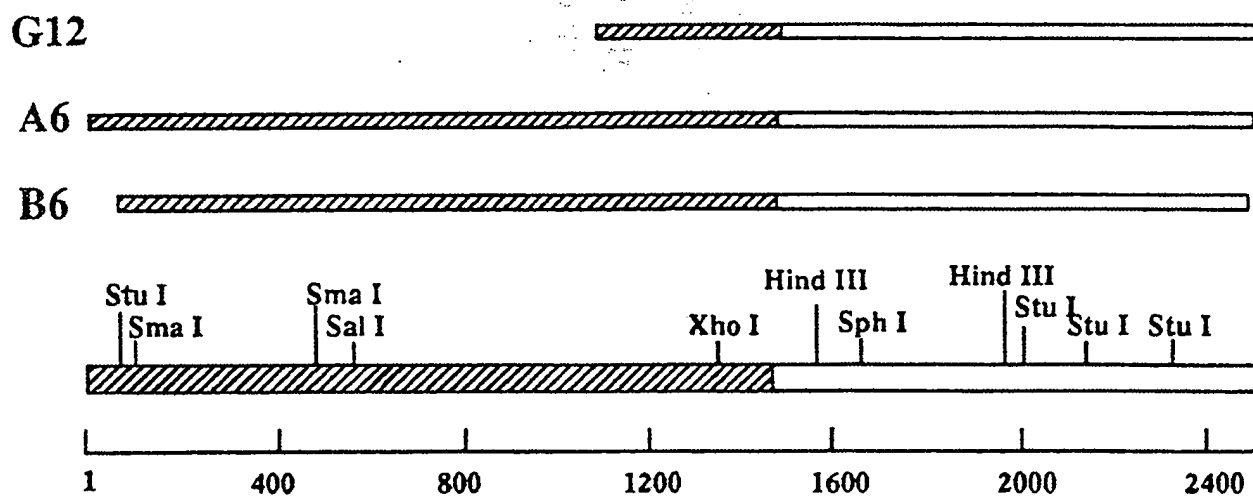


FIG. 4A

5/17

CTTTGCAGGCGGGGGGGGAGCGGGGATCGAGCCCTCGCCGAGGCCCTGCCGCCCATGGGCCCGGCCGCGCCGCGCTGTC 90
L C R Q R R P G A E R D R A L A E A C R H G P A P P P P V
ACCGGGCGCGGGCCGTGAGCGTCATGGCCCTTGCCCGGGGCCCTGCGGGCGGCCCATGCGCGCCGGCGCTGGAGGCCCTGCTCGGG 180
T R A A R A V S V M A L A G A P A G G P C A P A L E A L L G
3CGGGCGGCTGCGGCTGCTCGACTCCTCGCAGATCGTCATCATCTCGCCGCGCAGGACGCCGCCCGCCGCGCTCCACCGGCCCC 270
A G A L R L L D S S Q I V I I S A A Q D A S A P P A P T G P
3CGGGCGCGCGGCCCTGCGACCCCTGACCTGCTGCTCTTCGCCACACCGCAGGCGGCCCGGCCACACCCAGTGCGCCGCGGCC 360
A A P A A G P C D P D L L L F A T P Q A P R P T P S A P R P
5CGCTCGGCCCGCGGTGAAGCGGAGGCTGGACCTGGAACTGACCATCAGTACCTGGCCGAGAGCAGTGGGCCAGCTCGGGGCAGA 450
A L G R P P V K R R L D L E T D H Q Y L A E S S G P A R G R
GGCCGCCATCCAGGAAAGGTGTGAATCCCCGGGGAGAAGTCACGCTATGAGACCTCACTGAATCTGACCACCAAGCGCTTCCTGGAG 540
G R H P G K G V K S P G E K S R Y E T S L N L T T K R F L E
CTGCTGAGCCACTCGGCTGACGGTGTGTCGACCTGAACCTGGGCTGCCGAGGTGCTGAAGTGCAGAGCGGCGCATCTATGACATCACC 630
L L S H S A D G V V D [] N W A A E V [] K V Q K R R [] Y D I T
AACGTCTTGAGGGCATCCAGCTCATTGCCAAGAAGTCCAAGAACCAATCCAGTCCAGTGGCGAGCCACACACAGTGGCGCTCGGCGGA 720
N V [] E G I Q L I A K S K N H I Q W L G S H T T V G V G G
CGGCTTGAGGGTTGACCCAGGACCTCCGACAGCTGCAGGAGCGCAGCAGCTGGACCACCTGATGAATATCTGTACTACGCAGCTG 810
R L E G L T Q D L R Q L Q E S E Q Q L D H L M N I C T T Q L
CGCTGCTCTCCGAGGACACTGACAGCCAGCGCCTGGCCTACGTGACGTGTCAGGACCTTCGTAGCATTTGCAGACCCCTGCAGAGCAGATG 900
R L L S E D T D S Q R L A Y V T C Q D L R S I A D P A E Q M
GTTATGTGATCAAAGCCCTCCTGAGACCCAGCTCCAAGCCGTGGACTCTTCGAGAACTTTCAGATCTCCCTTAAGAGCAAACAAGGC 990
V M V I K A P P E T Q L Q A V D S S E N F Q I S L K S K Q G

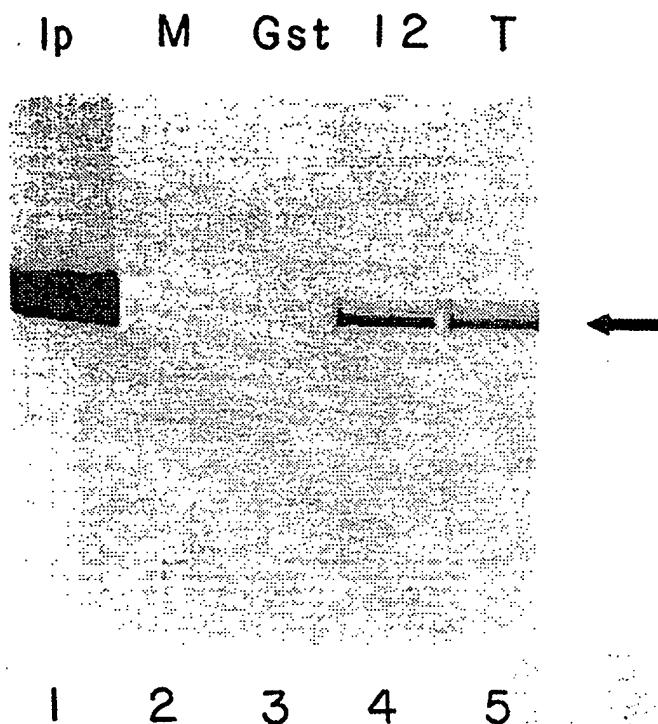
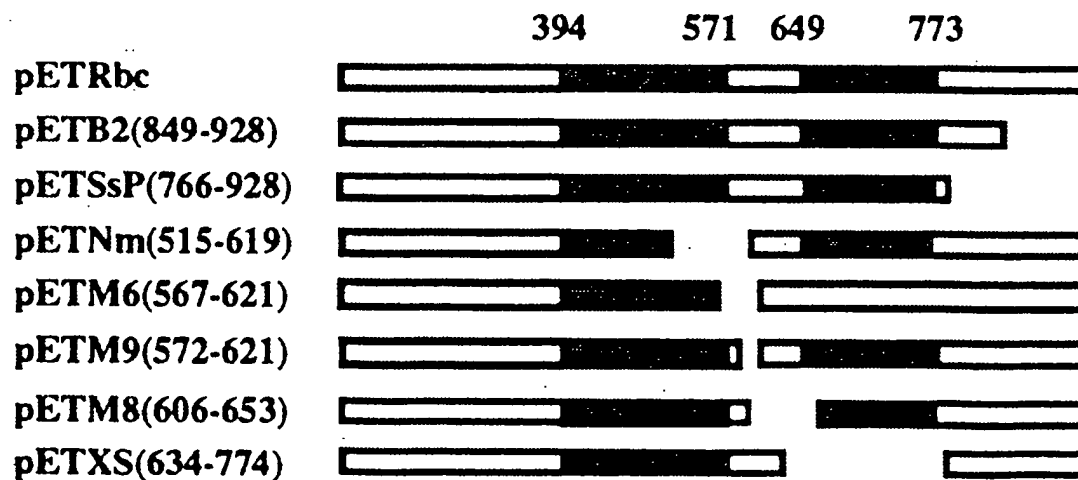
FIG. 4B-1

CCGATCGATGTTTTCTGTGCCCTGAGGAGACCGTAGGTGGGATCAGCCCTGGGAAGACCCCATCCAGGAGGTCAATTCTGAGGAGGAG 1080
P I D V F L C P E E T V G G I S P G K T P S Q E V T S E E E
AACAGGGCCACTGACTCTGCCACCACCATAGTGTCAACCACCACCATCATCTCCCCCTCATCCCTCACCACAGATCCCAGGCCAGTCTCTACTC 1170
N R A T D S A T I V S P P S S P P S S L T T D P S Q S L L
AGCCTGGAGCAAGACCGCTGTGTCCCGGATGGGCAGCCTGCGGGCTCCCGTGGACGAGGACCGCTGTCCCCGCTGGTGGCGGCCGAC 1260
S L E Q E P L L S R M G S L R A P V D E D R L S P L V A A D
TCGCTCCTGGAGCATGTGCGGGAGGACTTCTCCGGCCCTCCTCCCTGAGGAGTTCATCAGCCTTTCCCCACCCACGAGGCCCTCGACTAC 1350
S L L E H V R E D F S G L L P E E F I S L S P P H E A L D Y
CACTTCGGCCTCGAGGAGGCGAGGGCATCAGAGACCTCTTCGACTGTGACTTTGGGGACCTCACCCCTGGATTTCTGACAGGGCTTG 1440
H F G L E E G E G I R D L F D C D F G D L T P L D F
GAGGGACCAAGGTTTCCAGAGATGCTCACCTTGCTCTGCAGCCCTGGAGCCCTGTCCCTGGCCGTCTCCAGCCTGTTTGGAAACA 1530
TTTAATTTATACCCCTCTCTGTCTCCAGAGCTTCTAGCTCTGGGGCTGGCTACCGCTAGGAGGCTGAGCAAGCCAGGAAGGAAG
GAGTCTGTGTGTGTATGTGCATGCAGCCTACACCCACACGTGTGTACGGGGGTGAATGTGTGTGAGCATGTGTGTGCATGTGTGTACC
GGGGAATGAAGTGAACATACACCTCTGTGTGCATGCAGACACGCCCCAGTGTGTCCACATGTGTGTGCATGAGTCCATGTGTGCCG
GTGGGGGGCTCTAAGTGCATCTTCGGCCCTTTTGGCTGGCGTGTAGGACGGTGAGACACTTCTGTCTTAAGGTTTTTCTGATTGAAGCTTT
GACCAGGCCAGTGGGAGGCTTTGGCTGGCTGGCGTGTAGGACGGCTGGGGAATCAGCAAGGGGAGGAGGGGTGTGGGGTTGATACCCCAACTC
AATGGAGCGTTATTTATCGAGGCCCTCTTTGGTGAGCCTGGGGAATCAGCAAGGGGAGGAGGGGTGTGGGGTTGATACCCCAACTC
CCTCTACCCCTTGAGCAAGGGCAGGGTCCCTGAGCTGTTCTTCTGCCCATACTGAGGAACTGAGGCCCTGGGTGATTTATTTATTGGGA
AAGTGAGGGAGGAGACAGACTGACTGACAGCCATGGGTGGTCAGATGGTGGGTGGGCCCTCTCCAGGGGGCCAGTTTCCAGGGCCCCAGC
TGCCCCCAGGATGGATATGAGATGGGAGAGGTGAGTGGGGACCTTCACTGATGTGGGCAGGAGGGGTGTGAAGGCCTCCCCCAGCCCC
AGACCCCTGTGTCCTCTGACGTGCTGAAGCGCCTGCCCTCCCACTGCTGTGCCCACTCCCAATCTGCACTTTGATTGCTTTCCTA
ACAGCTCTGTTCCCTCTGCTTTGGTTTTTAATAAATATTTTGATGACGTTAAAAAATAAAA 2492

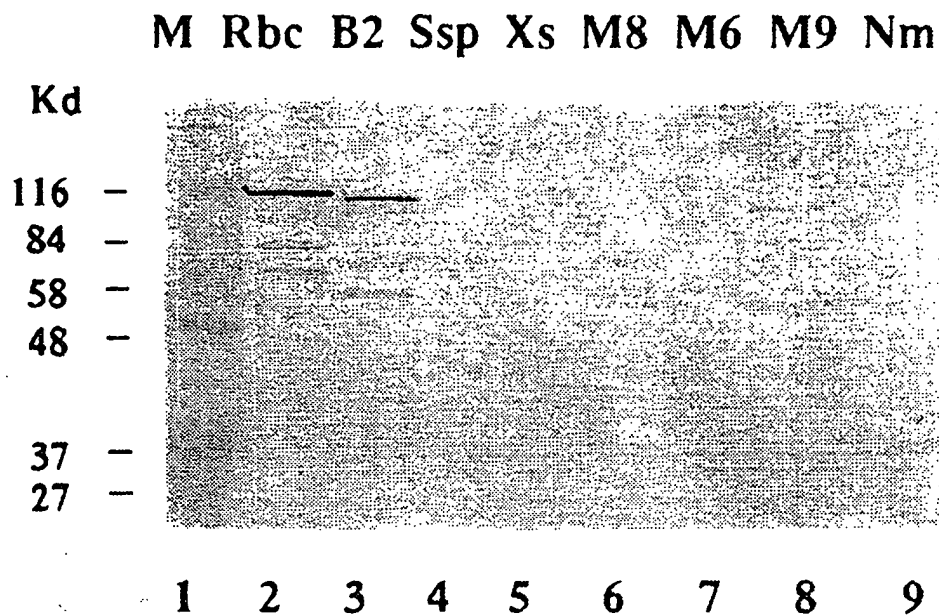
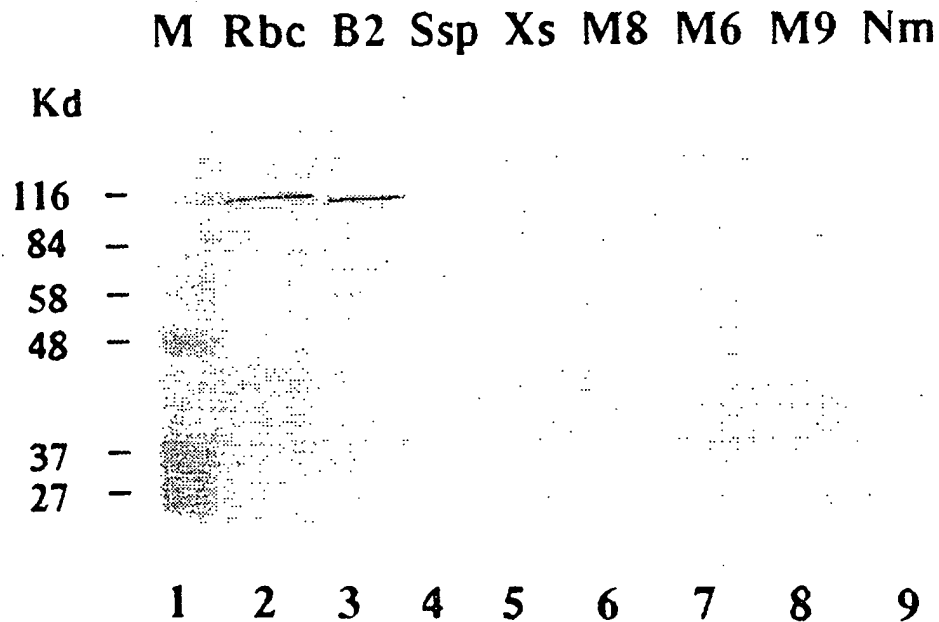
6 / 17

FIG. 4B-2

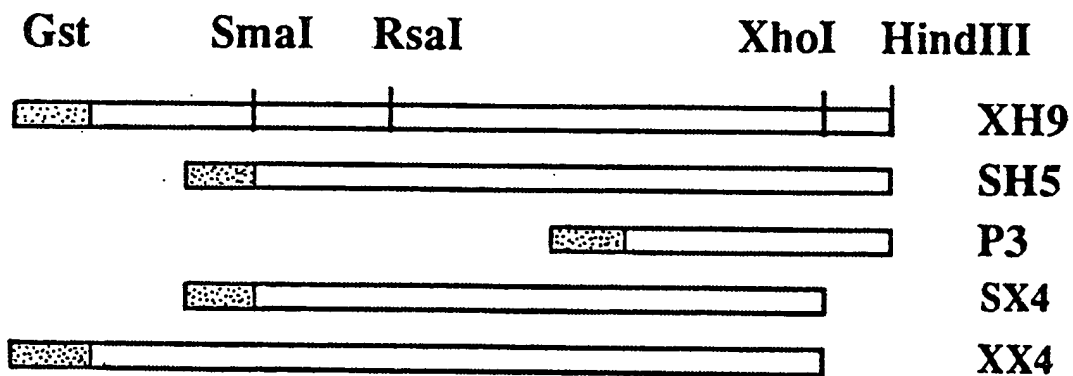
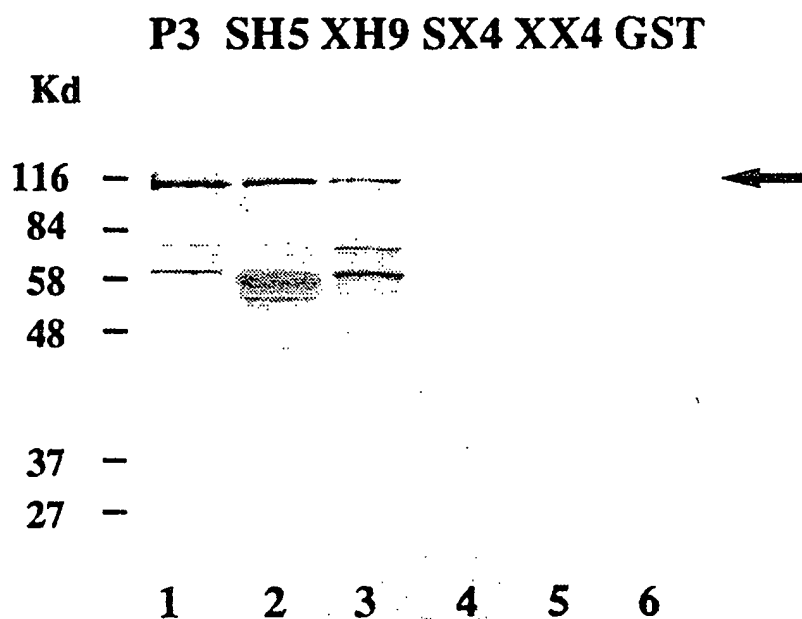
7/17

FIG. 5A**FIG. 5B****SUBSTITUTEHEET**

8/17

FIG. 5C**FIG. 5D****SUBSTITUTEHEET**

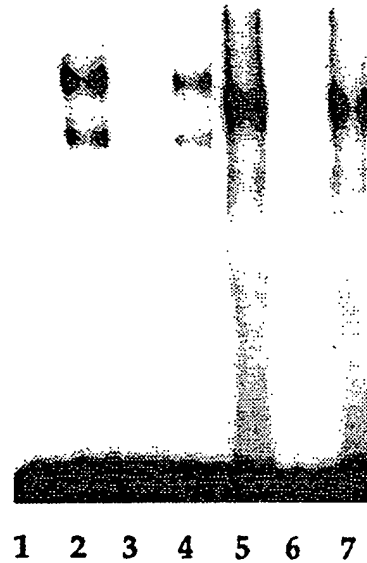
9/17

FIG. 6A**FIG. 6B**

10/17

E2F	-	+	+	+	-	-	-
SH5	-	-	-	-	+	+	+
c mp.	-	-	W	M	-	W	M

FIG. 7A



RB	-	-	+	+	-	+
SH5	-	+	+	-	+	+

FIG. 7B



SUBSTITUTE SHEET

11/17

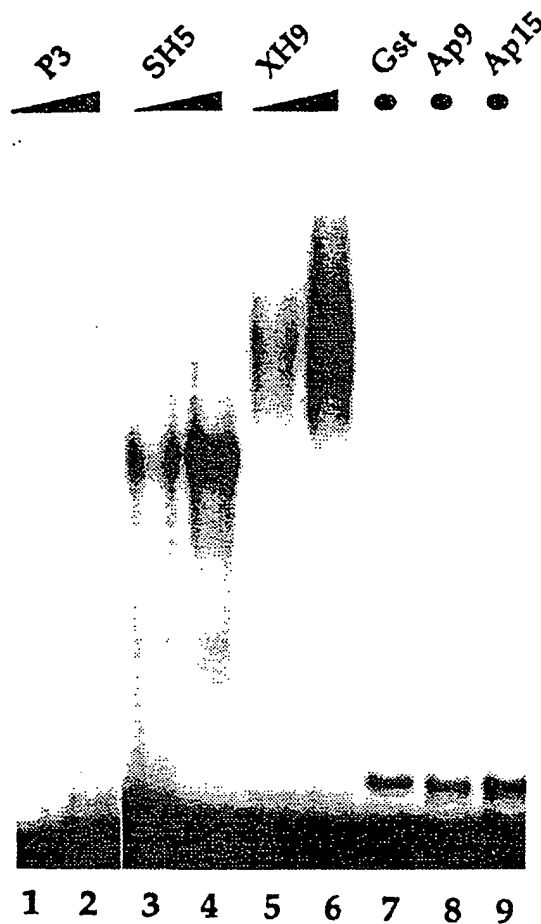


FIG. 7C

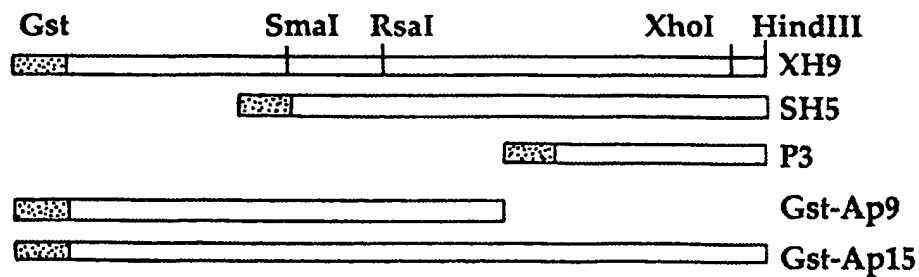


FIG. 7D

UBSTITUTE SHEET

12 / 17

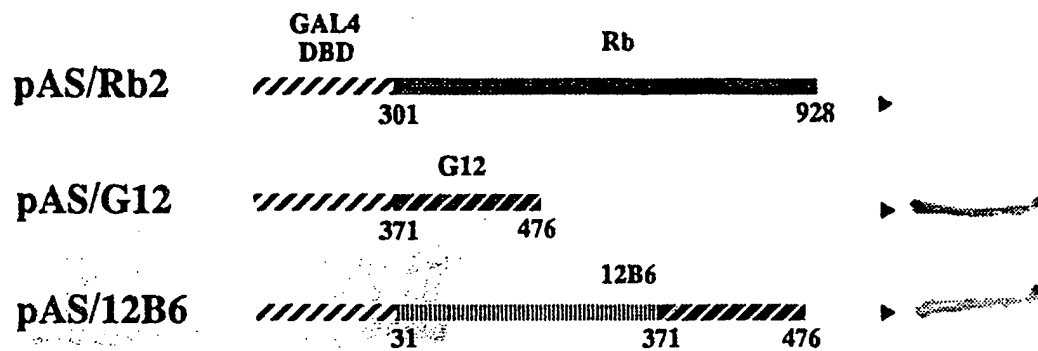


FIG. 8

13/17

FIG. 9A

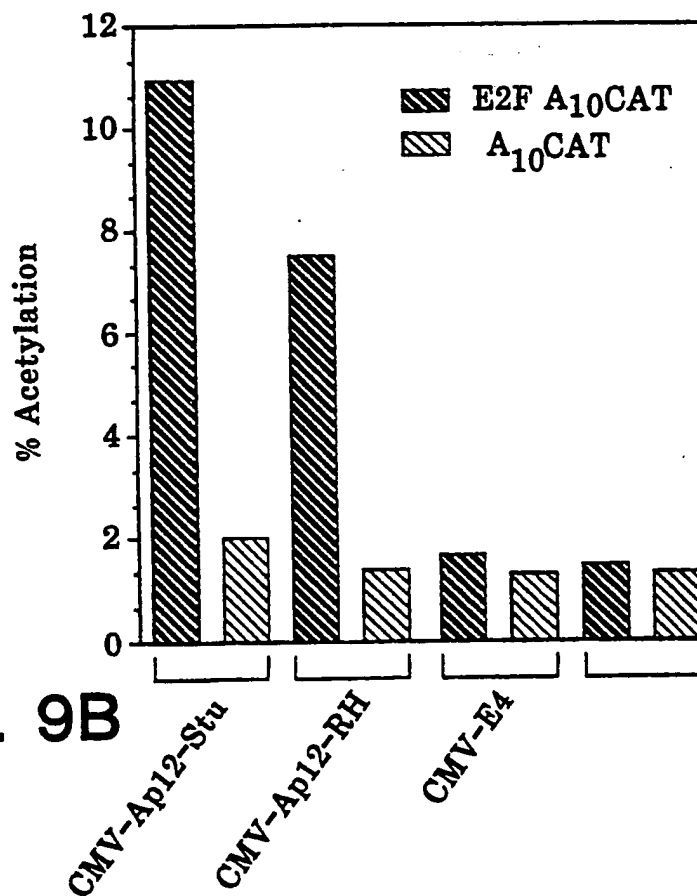
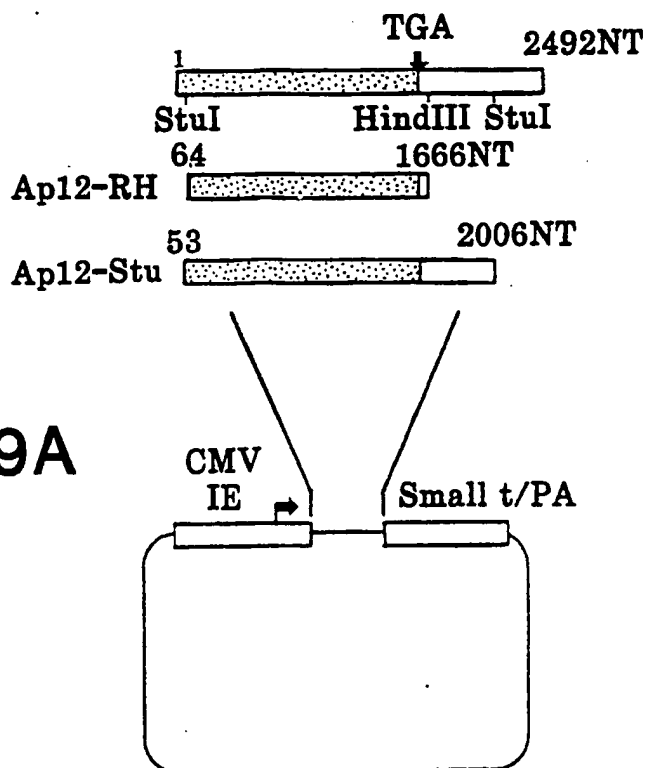


FIG. 9B

SUBSTITUTE SHEET

14/17

KbAp2p

10 20 30 40 50 60 70 80
 CGCCTTGACCTTGCTGGGAATGCTCGGTCAGACAAGGGCAGCATGTCTGAAGACTGTGGGCCAGGAACCTCCGGGGAGCT 80
 GGGCGGCTGAGGCGATCAAAATTGAGCCAGAGGATCTGGACATCATTGAGGTACCGTCCAGACCCCTCGCCAACCTCT 160
 GAGGAAATGACAGACTCG 178

RbAp2r

10 20 30 40 50 60 70 80
 TTTTACTTATTTAAAAAGGCCTTGGTGGCAGGAATATAGTGTAATAATCATTGGAAAACTAAAAGGCATCGATACAT 80
 ATCCGAATATACATTTTGACATAAATTACATTTCTTTAGTCTTTCTGAGTGAGGTCTGATTCACTACT 151

FIG. 10

RbAp8p

10 20 30 40 50 60 70 80
 TTTACGACAGAGCACTATTGCCAAGCGTTCAAATGCAGCACCATTAAGTAACACAAAAAAGCATCTGGGAAGACTGTAT 80
 CTACTGCTAAAGCAGGAGTGAAACAACCAGAAAGGAGTCAGGTAAAGAAGAAGTATGTATGCTACTGAAACCTGAGTAC 160
 CATAAGGAGAATAGAAGGTGCAGCCGAAATAGCGGACAAATTGAAGTGGATACCTGAAGTATCAGTGTCTTCAAGTCATT 240
 CTTCACTGTCTATCTT 255

RbAp8r

10 20 30 40 50 60 70 80
 GAATTCAACTGTAGCTTGGTTTTCAAAGTATCTGGATCTAGTATTTTCACTCTTTTGTCTTCTTCAGCACAAACATTTTA 80
 CACAGACATATTCTTTGTCTTCTCGCCATCTGCTGTGCTTGAGAAAGACTTAACCCAACACAATCACCATGAAACCGAG 160
 TCATCACATCTCCACAGCCAACCATAACTGTTGCATGTGTTTTGCAAACCACTGTTGCTGGAGTCACATATATTCGT 240
 TCAAT 245

FIG. 11

RbAp15p

10 20 30 40 50 60 70 80
 GAATTCACTGGAGCACCAGTAGAAGGTGCAGGAGAAGAGGCATTGACTCCATCAGTTCCTATAATAAAGGTCCCAAACC 80
 TAAGAGGGGAGAAGAAGGAGCCTGGTACCAGAGTGAGAAAAACACCTACATCATCTGGTAAACCTAGTGCAAAGAAAGTGA 160
 AGAAACGGAATCCTTGGTCAGATGATGAATCCAAGTCAGAAAGTGATTTGGAAGAAACAGAACCTGTGGTTATTCCAAGA 240
 GATTCTTTGCTTAGGAGAGCAGCAGCCGAAAGACCTAAATACACATTTAATTTCTCAGAAGAAGAGGATGATGATGCTGA 320
 TGATGATGATGATGACAATAATGATTTAGAGGAATTGAAAGTTAAAGCATCTCCATAACAAATGATGGGGAAGATGAAT 400
 410 420 430 440 450 460 470 480
 TTGTTCTTTCAGATGGGTTAGATAAAGATGAATATACATTTTACCAGGCAAATCAAAGCCTCACCAGAAAAATCTTTG 480
 CATGACAAAAAAGTCAGGATTTTGGAAATCTTCTCATTTCTTCATATTCTCAGAAGTCAGAAGATGATTGAGCTAA 560
 ATTTGACAGTAATGAAGAAGATTCTGCTTCTGTTTTTACCATCATTTGGTCTGAAACAGACAGATAAAGTTCCAAGTA 640
 AAACGGTAGCTGCTAAAAAGGGGAAACCGTCTTCAGATACAGTCCCTA 688

RbAp15r

10 20 30 40 50 60 70 80
 GCAATGTTTAATTAAGTGGGGAAAGAGCACAAACATTTTTCAACAAATACTTGTTGTTGCTTTTTGTCTTCTCTGTCTCA 80
 GACCTTTTGTACATCTGGCTTATTTAATGTGATGATGAATTGACCGTTTTTTATTATTGTGGTAGGCCCTTTTAACATT 160
 TTGTTCTTACACATACAGTTTTATGCTCTTTTTTACTCATTGAAATGTACAGTACTGTCTGATTGGCTTTGTAGAATTGGT 240
 TATAGACTGCCGTGCATTAGCACAGATTTTAATTGTGTCATGGTTACAAACTACAGACCTGCTTTTTGAAATGAAATTTAA 320
 CATTAAAAATGGAAGTGTGAAAAA 348

FIG. 12

SUBSTITUTE SHEET

15/17

RbAp4

10 20 30 40 50 60 70 80
GAATTCGGGCCAAGAAGCCTAATGAGAAAAACAAACCACTTGATAATAAGGGAGAAAAAGAAAAAGAAAACTGAAGA 80
AAAAGGCGTAGATAAAGATTTTGAGTCTTCTTCAATGAAAATCTCGAACTAGAAGTGACTGAAATAGTGAAACCATCAC 160
CAAAGCGCAAAATGGAACCTGATACTGAAAAATGGATAGGACCCCTGAAAAGGACAAAATTTCTTTAAGTGCGCCAGCC 240
AAAAAATCAAACCTCAACAGAGAACTGGGAAGAAAATTGGAAGTACAGAAAATATATCAAACACAAAAGAACCCTCTGA 320
AAAATTGGAGTCAACATCTAGCAAAGTTAAACAAGAAAAAGTCAAAGGAAAGGTCAGACGAAAAGTGACTGGAAGTGAAG 400
410 420 430 440 450 460 470 480
GATCCAGCTCAACTCTGGTGGATTACaCCaGTACGAGCTCAACTGGAGGCAGTCCTGTGCGGAAATCTGAAGAAAAACA 480
GATACAAAGCGAACTGTGATTAAACGATGGAAGAATAATAATGACAATACCGCGCCACGTGAAGATGTTATCATTAT 560
GATTCAAGTTCTCTCAATCCAAATGGGATAAAGATGACTTTGAATCTGAAGAAGAAGATGTTAAATCCACACAGCCTATAT 640
CAAGTGTAGGAAAACCTGCTAGTGTATAAAAAATGTTAGTACAAAGCCATCAAATATAGTCAAGTATCCTGAGAAAGAA 720
AGTGAGCCATCCGAGAAAATTGAGAAATTCACCAAGGACGTGAGCCATGAAATCATACAACATGAGGTTAAAGTTCAAA 800
810 820 830 840 850 860 870 880
AAACTCTGCATCTAGTGAAAAAGGGAAAAACAAAGATCGAGATTATTCAGTGTTGGAAAAGGAGAACCTGAAAAGAGGA 880
AGAACAGCACTCAGCCAGAGAAAGAGAGTAATTTGGACCGTCTGAATGAACAAGGAAATTTTAAAGTCTGTCTCAATCT 960
TCCAAAGAGGCTAGAACGTCAGATAAACATGATTCCACTCGTGCTTCTCAATAAAGACTTCACTCCCAATAGAGACAA 1040
AAAACTGACTATGACACCAGAGAGTATTCAAGTTCCAAAcgTAGAGATGAAAAGAATGAATTAACAAGACGAAAAGACT 1120
CTCCTTCTCGGAATAAAGATTCTGCATCTGGACAGAAAAATAAACCAAGGGAAGAGAGAGATTTGCCTAAAAAGGAACA 1200
1210 1220 1230 1240 1250 1260 1270 1280
GGAGATTCCAAAAAAGTAATTCTAGTCCCTCAAGAGACAGAAAACCTCATGATCACAAGCCACTTATGATACTAAACG 1280
GCCAAATGAAGAGACAAAATCTGTAGATAAAAAATCCTTGTAGGATCGTGAGAAGCATGTATTAGAAGCAAGGAACAATA 1360
AAGAGTCAAGTGGCAATAAAcTqCTTTATATACTTAACCCACCAGAGAcAcAGGTTGAAAAGAGCAAATTAAGTGGGCAA 1440
ATTGACAAGAGTACTGTCAAGCCTAAACCCAGTTAAGTCATTCTCTAGACTTTCTCTGACTTAAGTGAAGAACTCA 1520
TGAAGCTGCTTTTGAACCAGACTATAATGAAAGTGACAGTGAAAGTAATGTTTCTGTAAAAGAAGAGGAATCTTCAGGAA 1600
1610 1620 1630 1640 1650 1660 1670 1680
ACATTTCTAAGGACCTGAAAGATAAAATAGTGGAGAAAGCAAAAGAGAGCCTGGACACAGCAGCAGTTGTCCAGGTGGGC 1680
ATAAGCAGGAATCAGAGCCACAGCAGCCCCAGCGTCAGCCCCAGCAGAGCCACAGTCCTTCTGGAAGCCAGACCCGAAG 1760
CCACAGTAGCAGTGCCAGCTCAGCAGAAAGTCAGGACAGC 1800

FIG. 13

16 / 17

RbAp10

10 20 30 40 50 60 70 80
GAATTCGGCCGGAATTAATTCGGGGATTTCCTGGGGAATCAGGAAGATATCCATAATCTTCAACTGCGGGTAAAAGAG 80
ACATCAAATGAGAATTTGAGATTACTTCATGTGATAGAGGACCGTGACAGAAAAGTTGAAAGTTTGCTAAATGAAATGAA 160
AGAATTAGACTCAAACTCCATTTACAGGAGGTACAATAATGACCAAAATTGAAGCATGCATAGAATTGGAAAAAATAG 240
TTGGGGAACCTAAGAAAGAAAACCTCAGATTTAAGTGAAAAATTGGAATATTTTTCTTGATCACCAGGAGTTACTCCAG 320
AGAGTAGAACTTCTGAAGGCCTCAATTCTGATTTAGAAATGCATGCAGATAAATCATCACGTGAAGATATTGGAGATAA 400
410 420 430 440 450 460 470 480
TGTGGCCAAGGTGAATGACAGCTGGAAGGAGAGATTTCTTGATGTGGAAAATGAGCTGAGTAGGATCAGATCGGAGAAAAG 480
CTAGCATTTGAGCATGAAGCCCTCTACCTGGAGGCTGACTTAGAGGTAGTTCAAACAGAGAAGCTATGTTTAGAAAAAGAC 560
AATGAAAATAAGCAGAGAAGGTTATTGTCTGCCTTGAAGAAGAACTCTCAGTGGTCACAAGTGAGAGAAACCAGCTTCGTGG 640
AGAATTAGATACTATGTCAAAAAAACACGGCACTGGATCAGTTGCTGAAAAAATGAAGGAGAAAACACAAGAGCTTG 720
AGTCTCATCAAAGTGAGTGTCTCCATTGCATTAGGTGGCAGAGGCAGAGGTGAAGGAAAAGACGGAACCTCTTCAGACT 800
810 820 830 840 850 860 870 880
TTGTCTCTGATGTGAGTGAGCTGTTAAAAGACAAAACCTCATCTCCAGGAAAAGCTGCAGAGTTTGGAAAAGGACTCACA 880
GGCACTGTCTTTGACAAAATGTGAGCTGGAAAACCAAATGCACAACCTGAATAAAGAGAAAAGAATTGCTTGCAAGGAAT 960
CTGAAAGCCTGCAGGCCAGACTGAGTGAATCAGATTATGAAAAGCTGAATGTCTCCAAGGCTTGGAGGCCGCACTGGTG 1040
GAGAAAGGTGAGTTCGCATTGAGGCTGAGCTCAACACAGGAGGAAGTGCATCAGCTGAGAAGAGGCATCGAGAACTGAG 1120
AGTTCCGATTGAGGCCGATGAAAAGAAGCAGCTGCACATCGCAGAGAACTGAAAGAACGCGAGCGGGAGAATGATTCAC 1200
1210 1220 1230 1240 1250 1260 1270 1280
TTAAGGATAAAGTTGAGAACCTTGAAAGGGAATTGCAGATGTCAGAAGAAAACAGGAGCTAGTGATTCTTGATGCCGAG 1280
AATTCCAAAGCAGAAGTAGAGACTCTAAAAACCAAATAGAAGAGATGGCCAGAAGCCTGAAAGTTTTGAATTAGACCT 1360
TGTCACGTTAAGGTCTGAAAAAGAAAATCTGACAAAACAAATACAAGAAAACAAGGTCAAGTTGTCAGAACTAGACAAGT 1440
TACTCTCTTCATTTAAAAGTCTGTTAGAAGAAAAGGAGCAAGCAGAGATACAGATCAAAGAAGAATCTAAAACCTGCAGTG 1520
GAGATGCTTCAGAATCAGTTAAAGGAGCTAAATGAGGCAGTAGCAGCCTTGTTGTTGACCAAGAAATATGAAGGCCAC 1600
1610 1620 1630 1640 1650 1660 1670 1680
AGAACAGAGTCTAGACCCACCAATAGAGGAAGAGCATCAGCTGAGAAATAGCATTGAAAAGCTGAGAGCCCGCTAGAAG 1680
CTGATGAAAAGAGCAGCTCTGTGTCTTACAACAACTGAAGGAAAGTGAGCATCATGCAGATTTACTTAAGGGTAGAGTG 1760
GAGAACCTTGAAAGAGAGCTAGAGATAGCCAGGACAAACCAAGAGCATGCAGCTCTTGAGGCAGAGAATCCAAAGGAGA 1840
GGTAGAGACCCTAAAAGCAAAAATAGAAGGGATGACCCAAAGTCTGAGAGGTCTGGAATTAGATGTTGTACTATAAGGT 1920
CAGAAAAAGAAAATCTGACAAATGAATTACAAAAGAGCAAGAGCGAATATCTGAATTAGAAATAATAAATTCATCATTT 2000
2010 2020 2030 2040 2050 2060 2070 2080
GAAAATATTTTGAAGAAAAAGAGCAAGAGAAAGTACAGATGAAAGAAAAATCAAGCACTGCCATGGAGATGCTTCAAAC 2080
ACAATTAAGAGAGCTCAATGAGAGAGTGGCAGCCCTGCATAATGACCAAGAAGCCTGTAAGGCCAAAGAGCAGAATCTTA 2160
GTAGTCAAGTAGAGTGTCTTGAACCTTGAGAAGGCTCAGTTGCTACAAGGCCTTGATGAGGCCAAAAATAATTATATTGTT 2240
TTGCAATCTTCAGTGAATGGCCTCATTCAAGAAAGTGAAGATGGCAAGCAGAAAACCTGGAGAAGAAGGATGAAGAAATCAG 2320
TAGACTGAAAAATCAAATTCAGACCAAGAGCAGCTTGCTCTAACTGTCCAGGTGGAAGGAGAGCACCAACTTTGGA 2400
2410 2420 2430 2440 2450 2460 2470 2480
AGGAGCAAACTTAGAACTGAGAAATCTGACAGTGAATTTGGAGCAGAAGATCCAAGTGCTACAATCCAAAAATGCCTCT 2480
TTGCAGGACACATTAGAAGTGCTGCAGAGTTCTTACAAGAATCTAGAGAATGAGCTTGAATTGACAAAAATGGACAAAAT 2560
GTCCTTTGTTGAAAAAGTAAACAAAATGACTGCAAGGAAACTGAGCTGCAGAGGGAATGCATGAGATGGCACAGAAAA 2640
CAGCAGAGCTGCAAGAAGAACTCAGTGGAGAGAAAAATAGGCTAGCTGGAGAGTTGCAGTTACTGTTGGAAGAAATAAAG 2720
AGCAGCAAAGATCAATTGAAGGAGCTCACACTAGAAAAATAGTGAATTGAAGAAGAGCCTAGATTGCATGCACAAAGACCA 2800

FIG. 14A

SUBSTITUTE SHEET

17/17

RbAp10 CONT.

2810 2820 2830 2840 2850 2860 2870 2880
GGTGGAAAAGGAAGGGAAAGTGAGAGAGGAAATAGCTGAATATCAGCTACGGCTTCATGAAGCTGAAAAGAAACACCAGG 2880
CTTTGCTTTTGGACACAAACAAACAGTATGAAGTAGAAATCCAGACATACCGAGAGAAATTGACTTCTAAAGAAGAATGT 2960
CTCAGTTCACAGAAGCTGGAGATAGACCTTTTAAAGTCTAGTAAAGAAGAGCTCAATAATTCATTGAAAGCTACTACTCA 3040
GATTTTGGAGAATTTGAAGAAAACCAAGATGGACAATCTAAAATATGTAAATCAGTTGAAGAAGGAAAATGAACGTGCCC 3120
AGGGGAAAATGAAGTTGTTGATCAAATCCTGTAAACAGCTGGAAGAGGAAAAGGAGATACTGCAGAAAGAACTCTCTCAA 3200
3210 3220 3230 3240 3250 3260 3270 3280
CTTCAAGCTGCACAGGAGAAGCAGAAAACAGGTAAGTGTATGGATACCAAGGTCGATGAATTAACAAGTGAAGTCAAGA 3280
ACTGAAAGAACTCTTGAAGAAAAACCAAGGAGGCAGATGAATACTTGGATAAGTACTGTTCTTGTATTATAAGCCATG 3360
AAAAGTTAGAGAAAGCTAAAGAGATGTTAGAGACACAAGTGGCCCATCTGTGTTTACAGCAATCTAAACAAGATTCCCGA 3440
GGGTCTCCTTTGCTAGGTCCAGTTGTTCCAGGACCATCTCCAATCCCTTCTGTTACTGAAAAGAGGTTATCATCTGGCCA 3520
AAATAAGCTTCAGGCAAGAGGCAAGATCCAGTGAATATGGGAGAATGGTGGAGGACCAACACCTGCTACCCAGAGA 3600
3610 3620 3630 3640 3650 3660 3670 3680
GCTTTTCTAAAAAAGCAAGAAAGCAGTCATGAGTGGTATTACCCCTGCAGAAGACACGGAAGGTAAGTGTGAGTTGAGCCA 3680
GAGGGACTTCCAGAAGTTGTAAAGAAAGGGTTTGTGACATCCCGACAGGAAAGACTAGCCCATATATCTGCGAAGAAC 3760
AACCATGGCAACTCGGACCAGCCCCCGCTGGCTGCACAGAAGTTAGCGCTATCCCGACTGAGTCTCGGCAAGAAAAATC 3840
TTGCAGAGTCTCCAAACCAACAGCTGGTGGCAGCAGATCACAAAAGGTCAAAGTTGCTCAGCGGAGCCCAGTAGATTCA 3920
GGCACCATCCTCCGAGAACCACACGAAATCCGTCCAGTCAATAATCTTCTGAGAGAAGTCCGACTGACAGCCCCAG 4000
4010 4020 4030 4040 4050 4060 4070 4080
AGAGGGCTTGAGGGTCAAGCGCCGCGACTTGTCCCCAGCCCCAAAGCTGGACTGGAGTCCAAGGGCAGTGAGAACTGTA 4080
AGGTCCAGTGAAGGCACTTTGTGTGTCAGTACCCCTGGGAGGTGCCAGTCATTGAATAGATAAGGCTGTGCCTACAGGAC 4160
TTCTCTTTAGTCAGGGCATGCTTTATTAGTGAGGAGAAAAACAATTCCTTAGAAGTCTTAAATATATTGTACTCTTTAGAT 4240
CTCCCATGTGTAGGTATTGAAAAAGTTTGGAAAGCACTGATCACCTGTTAGCATTGCCATTCTCTACTGCAATGTAAATA 4320
GTATAAGCTATGTATATAAAGCTTTTGGTAATATGTTACAATTAATGACAAGCACTATATACAATCTCTGTTTGT 4400
4410 4420 4430 4440 4450 4460 4470 4480
ATGTGGGTTTTACTATAAAAAATGCAAAACACATTTTATTCTTCTAATTAACAGCTCCTAGGAAAATGTAGACTTTTGC 4480
TTTATGATATTCTATCTGTAGTATGAGGCATGGAATAGTTTTGTATCGGGAATTTCTCAGAGCTGAGTAAATGAAGGAA 4560
AAGCATGTTATGTGTTTTTAAGGAAAATGTGCACACATATACATGTAGGAGTGTATCTTTCTTACAATCTGTTTTA 4640
GACATCTTTGCTTATGAAACCTGTACATATGTGTGTGGGTATGTGTTTATTTCCAGTGAGGGCTGCAGGCTTCCTAGA 4720
GGTGTGCTATACCATGCGTCTGTGCTTGTGCTTTTTTCTGTTTTTAGACCAATTTTTTACAGTTCTTTGGTAAGCATTGT 4800
4810 4820 4830 4840 4850 4860 4870 4880
CGTATCTGGTGATGGATTAACATATAGCCTTTGTTTTCTAATAAAATAGTCGCCTTCGTAAAAA 4868

FIG. 14B

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :Please See Extra Sheet.

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.1, 69.1, 172.3, 240.2, 320.1; 436/501, 536; 530/350, 388.15, 389.1; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, EMBASE

Search terms: retinoblastoma-associated polypeptide, retinoblastoma binding, E2F factor, transcription factor

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Cell, Volume 64, issued 08 February 1991, W.G. Kaelin, Jr. et al, "Identification of cellular proteins that can interact specifically with the T/E1A-binding region of the retinoblastoma gene product", pages 521-532, see entire document.	6-8, 21
X	Nature, Volume 352, issued 18 July 1991, D. Defeo-Jones et al, "Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product", pages 251-254, see entire document.	1-13, 20-21
X	Cell, Volume 70, issued 24 July 1992, W.G. Kaelin, Jr. et al, "Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties", pages 351-364, see entire document.	1-21
X	Cell, Volume 70, issued 24 July 1992, K. Helin et al, "A cDNA encoding a pRB-binding protein with properties of the transcription factor E2F", pages 337-350, see entire document.	1-21

☒ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be part of particular relevance	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z*	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

18 FEBRUARY 1994

Date of mailing of the international search report

07 MAR 1994

 Name and mailing address of the ISA/US
 Commissioner of Patents and Trademarks
 Box PCT
 Washington, D.C. 20231

Facsimile No. NOT APPLICABLE

Authorized officer

DIAN C. JACOBSON

Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Molecular and Cellular Biology, Volume 12, Number 10, issued October 1992, S.K. Ray et al, "Identification of a 60-kilodalton Rb-binding protein, RBP60, that allows the Rb-E2F complex to bind DNA", pages 4327-4333, see entire document.	6-8, 21
Y	Science, Volume 258, issued 16 October 1992, J.R. Nevins, "E2F: A link between the Rb tumor suppressor protein and viral oncoproteins", pages 424-429, see pages 427-428.	1-13, 20-21
Y	Nature, Volume 352, issued 18 July 1991, L.R. Bandara et al, "Cyclin A and the retinoblastoma gene product complex with a common transcription factor", pages 249-251, see entire document.	6-8, 21
P,X	Biochemical and Biophysical Research Communications, Volume 194, Number 2, issued 30 July 1993, A.A. Ali et al, "Retinoblastoma gene product-associated proteins in human colon cancer cell lines", pages 848-854, see entire document.	1-21

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (5):

C07H 21/04; C07K 13/00, 15/28; C12N 5/00, 5/10, 15/00, 15/12, 15/85, 15/86; G01N 33/53

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/7.1, 69.1, 172.3, 240.2, 320.1; 436/501; 530/350, 388.15, 389.1; 536/23.5